REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE:	3. REPORT TYPE AND DATES COVERED
		Final Report 15-Aug-2005 - 30-Sep-2006

4. TITLE AND SUBTITLE

Co-Channel Interference Mitigation for Robust Coexistence of Frequency Hopped Networks

5. FUNDING NUMBERS

W911NF-05-1-0485

6. AUTHORS

Xiangqian Liu

8. PERFORMING ORGANIZATION REPORT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES

University of Louisville
Research Foundation, Inc.
2301 S. Third Street
Louisville, KY 40208 -

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

48740-CI.1

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The abstract is below since many authors do not follow the 200 word limit

14. SUBJECT TERMS

Frequency hopping, coexistence, co-channel interference, multiuser separation, hop timing estimation

15. NUMBER OF PAGES

Unknown due to possible attachments

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Report Title

Co-channel Interference Mitigation for Robust Coexistence of Frequency Hopped Networks

ABSTRACT

To meet the military's increased need for rapidly deployable communication solutions, wireless networks are becoming increasingly common within tactical environments. Frequency hopping (FH) is widely used for radio transmission in such networks due to its low probability of detection/interception. With the increasing deployment, multiple networks occupying overlapping frequency bands are likely to coexist in a physical environment, especially in tactical operations, emergency situations, or dense-populated areas. Consequently co-channel interference due to frequency collisions can become a major performance limiting factor. In this project, we developed a novel interference mitigation technique based on multidimensional frequency estimation coupled with the expectation-maximization principle, which effectively resolves collisions among non-collaborative networks, thus enabling robust coexistence of multiple FH networks. To deal with possible receiver-transmitter mismatch, we also designed a low-complexity model order variation detection method. Novel multidimensional frequency estimation algorithms are also investigated.

List of papers submitted or published that acknowledge ARO support during this reporting period. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

1. X. Liu, J. Li, and X. Ma, "An EM algorithm for blind hop timing estimation of multiple frequency hopped signals with bandwidth mismatch," IEEE Trans. Vehicular Technology, to appear, 2007.
2. J. Liu and X. Liu, "On the estimation of N-D frequencies from the eigenvectors of sample covariance matrix," IEEE Signal Processing Letters, to appear, Mar. 2007.
3. J. Liu and X. Liu, "An eigenvector-based approach for multidimensional frequency estimation with improved identifiability," IEEE Trans. Signal Processing, vol. 54, no. 12, pp. 4543--4556, Dec. 2006.
4. X. Liu, N. D. Sidiropoulos, and A. Swami, "Joint hop timing and frequency estimation for collision resolution in frequency-hopped networks," IEEE Trans. Wireless Communications, vol. 4, no. 6, pp. 3063--3074, Nov. 2005.

Number of Papers published in peer-reviewed journals: 4.00

(b) Papers published in non-peer-reviewed journals or in conference proceedings (N/A for none)

Number of Papers published in non peer-reviewed journals: 0.00

(c) Presentations

J. Li and X. Liu, "A software demonstration: collision resolution for robust coexistence of multiple wireless networks," The Fifth International Conference on Information Processing in Sensor Networks (IPSN), Nashville, TN, Apr. 19-21, 2006.

Number of Presentations: 1.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts): 0

Peer-Reviewed Conference Proceeding publications (other than abstracts):

1. J. Liu and X. Liu, "Joint 2-D DOA tracking for multiple moving targets using adaptive frequency estimation," Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), to appear, 2007.
2. J. Liu and X. Liu, "Time-varying channel identification and prediction in OFDM systems using 2-D frequency estimation," Proc. Military Communications Conference (MILCOM), Oct. 23-25, 2006.
3. J. Li and X. Liu, "A collision resolution technique for robust coexistence of multiple bluetooth piconets," The 64th IEEE Vehicular Technology Conference (VTC), Sept. 25-28, 2006.
4. J. Liu and X. Liu, "Optimizing eigenvector based frequency estimation in the presence of identical frequencies in multiple dimensions," Proc. the 7th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Cannes, France, Jul. 2-5, 2006.
5. J. Liu and X. Liu, "Statistical identifiability of multidimensional frequency estimation with finite snapshots," Proc. the 4th IEEE Workshop on Sensor Array and Multi-channel Processing (SAM), Waltham, MA, Jul. 12-14, 2006.
6. J. Li, X. Liu, and A. Swami, "Detection of model order variations in frequency hopped systems," in Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 680--684, New York City, NY, Jun. 2005.
7. J. Liu and X. Liu, "An eigenvector-based algebraic approach for two-dimensional frequency estimation with improved identifiability," in Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 655--659, New York City, NY, Jun. 2005.

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts): 7

(d) Manuscripts

1. J. Liu, X. Liu, and X. Ma, "Multidimensional frequency estimation with finite snapshots in the presence of identical frequencies," IEEE Trans. Signal Processing, submitted, May 2006.

Number of Manuscripts: 1.00

Number of Inventions:

Graduate Students

NAME	PERCENT_SUPPORTED	
Jun Liu	0.80	No
Jingli Li	0.20	No
FTE Equivalent:	**1.00**	
Total Number:	**2**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Xiangqian Liu	0.20	No
FTE Equivalent:	**0.20**	
Total Number:	**1**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Personnel receiving masters degrees

NAME

Total Number:

Names of personnel receiving PHDs

NAME

Total Number:

Names of other research staff

NAME PERCENT_SUPPORTED

FTE Equivalent:

Total Number:

Sub Contractors (DD882)

Inventions (DD882)

TECHNICAL REPORT

W911NF-05-1-0485

Co-channel Interference Mitigation for Robust Coexistence of Frequency Hopped Networks

August 15, 2005 – September 30, 2006

Xiangqian Liu, PI

Department of Electrical and Computer Engineering
University of Louisville, Louisville, KY 40292
Tel: 502-852-1559, Fax: 502-852-6807
Email: x.liu@louisville.edu

Contents

List of Figures

List of Tables

A Project Abstract

To meet the military's increased need for rapidly deployable communication solutions, wireless networks are becoming increasingly common within tactical environments. Frequency hopping (FH) is widely used for radio transmission in such networks due to its low probability of detection/interception. With the increasing deployment, multiple networks occupying overlapping frequency bands are likely to coexist in a physical environment, especially in tactical operations, emergency situations, or dense-populated areas. Consequently co-channel interference due to frequency collisions can become a major performance limiting factor. In this project, we developed a novel interference mitigation technique based on multidimensional frequency estimation coupled with the expectation-maximization principle, which effectively resolves collisions among non-collaborative networks, thus enabling robust coexistence of multiple FH networks. To deal with possible receiver-transmitter mismatch, we also designed a low-complexity model order variation detection method. Novel multidimensional frequency estimation algorithms are also investigated. The significance of this project in basic research lies in innovative schemes for collision resolution enabling interference-robust FH networking.

B Technical Report

B.1 Problem Statement

To meet the military's growing need for rapidly deployable communication solutions, wireless networks are becoming increasingly common in tactical environments. For example, military units (e.g., soldiers and tanks), equipped with wireless devices, could form multiple networks in tactical operations when they roam in a battlefield. Frequency hopping spread spectrum (FHSS) is widely used for radio transmission in such networks, due to its low probability of detection/interception [1]. For example, FHSS is adopted in the single channel ground-airborne radio system (SINCGARS), which is the current standard army combat net radio. Recently FHSS has also been adopted in commercial applications such as Bluetooth [2].

Because of their increasing deployment, multiple (homogeneous and/or heterogeneous) wireless networks with overlapping frequency bands are likely to coexist in a physical environment, especially in tactical operations, emergency situations, or dense-populated areas. Consequently co-channel interference due to frequency collisions can become a major performance limiting factor [3–5]. When collisions occur, network throughput decreases and delay can become excessive due to retransmissions. Theoretical analysis has shown that the packet error rate (PER) of one Bluetooth piconet due to collisions can be up to 10% if seven piconets coexist [6], and a Bluetooth receiver may experience up to 27% packet loss for data traffic in the presence of interference from an IEEE 802.11b based wireless local area network (WLAN) [7].

Recently the coexistence issue has gained increasing attention [8–14]. However, to date most coexistence schemes are designed for simultaneous functionality of a Bluetooth piconet and an 802.11b WLAN. The latter is a direct sequence spread spectrum (DSSS) network that occupies a fixed frequency band of 22 MHz. In this project, we study the problem of coexistence of multi-

ple FH networks, where the interfering frequency channels are constantly changing and the hop sequence of one network is unknown to another. We develop a physical layer method to mitigate co-channel interference and enable robust coexistence of multiple FH networks.

B.2 Prior Work

Most coexistence methods are developed for simultaneous function of a Bluetooth piconet and a WLAN, based on various MAC layer mechanisms that enable the Bluetooth devices to avoid hopping onto the frequency band occupied by the WLAN. Other methods mitigate co-channel interference with physical layer approaches.

B.2.1 Collision Avoidance for Coexistence of Bluetooth and WLAN

The collision avoidance techniques can be classified as collaborative and non-collaborative ones. For collaborative cases, attractive data transmission rates and throughput can be achieved by using a communication link between the Bluetooth and WLAN when they are embedded on the same device [15] or by coordinating the hop frequencies of the co-located Bluetooth devices [8].

Non-collaborative methods do not require direct communication between the two networks, and they usually rely on monitoring the channel to detect interference and estimate traffic. For example, power control is employed based on PER [9] or the received signal strength [10] to sustain the quality of service for a Bluetooth link. To avoid hopping onto preoccupied frequency channels, adaptive frequency hopping (AFH) modifies the Bluetooth frequency hopping sequence, and Bluetooth interference aware scheduling (BIAS) strategy postpones the transmission [11], both detecting preoccupied frequency bands by monitoring PERs on all channels. The overlap avoidance (OLA) schemes proposed in [12] are based on packet scheduling and traffic control. There are also hybrid methods that combine AFH and Bluetooth carrier sense (BCS) [13] or combine power control, listen-before-talk (LBT) and AFH [14] to achieve improved performance.

The center control mechanism needed for collaborative schemes [8, 15] confines their applications to certain situations. Power control methods [9, 10] depend on the accuracy of channel sensing and can not provide much improvement if the Bluetooth device is very close to the interfering device. Carrier sensing based schemes inevitably suffers from the hidden terminal problem [13, 16]. Approaches based on scheduling such as BIAS [11], OLA [12] and master delay MAC scheduling (MDMS) [17] cause delay in the transmission, hence they may not be bandwidth efficient. AFH [11] and interference source oriented AFH (ISOAFH) [17] are effective in dealing with WLAN interference, but not applicable for multiple co-located FH networks. The performance of AFH is also dependent on the update rate of the frequency classification to track the channel dynamics [11]. A hybrid method of power control, LBT and AFH proposed in [14] can achieve better performance but will add complexity to the application.

In summary, most non-collaborative interference detection and collision avoidance schemes are developed for coexistence of Bluetooth and WLAN, and they are not applicable to the coexistence of multiple FH networks, because the frequency channels are constantly changing and the hop sequence of one network is not known to another.

B.2.2 Co-channel Interference Mitigation in FH Systems

Instead of avoiding the collision by scheduling, many efforts have been made on interference suppression and mitigation on physical layer. The problem is challenging because when multiple FH wireless networks coexist in noncooperative scenarios, not only the hop sequences, hop timing/rate and other parameters such as hop bandwidth and bin-width are unknown to each other, but also there maybe model order variations even if the number of active emitters remains the same. If the receiver's bandwidth is not matched to the hop bandwidth of the emitters, the FH signals may hop in and out of the observation frequency band of the receiver.

Considerable work has been done on blind interference suppression for FHSS systems using an antenna array [18, 19]. These approaches aim at interference suppression rather than joint multiuser detection and hop timing estimation, and most of them require at least the knowledge of the hop sequence of the user of interest, and their interference nulling capability is bounded by the degrees of freedom in the adaptive array. Multiuser detection for FH systems have been considered in [20, 21]. They assume that the hop sequences of are known to the receiver, and thus they are not applicable to noncooperative FH emitters. In [22] the estimation of the hop sequence for a single user is discussed, while the remaining users are treated as white Gaussian interference. The approach of [22] is conceptually simple, but assumes perfect channel knowledge.

Without assuming the knowledge of hop sequences, several (semi-)blind methods have been proposed for hop timing and frequency estimation. For example, with known hop rate and frequency bins, channelized receivers have been proposed for semi-blind hop timing estimation [23, 24]. However, the performance of those receivers degrades rapidly if the channelization is imperfect, for example, when there is unknown bandwidth mismatch or carrier frequency offset. In [25, 26], hop timing and frequency estimators based on the principle of dynamic programming (DP) were developed for blind tracking of multiple FH signals, using either a decoupled approach [25] or a joint approach [26]. These methods neither assume knowledge of hop sequences or hop timing, nor rely on channelization, and hence are robust to frequency offset. However, they require accurate model order information, and thus can not handle bandwidth mismatch. In addition, the complexity of DP is too high to be feasible for practical implementation. An iterative maximum likelihood (ML) algorithm is developed in [27] to estimate the hop timing and frequencies with low complexity, but it can only deal with the single-user two-hop case (one hop timing instant to be estimated).

B.3 Summary of Major Results

In this project, we develop a physical layer method for co-channel interference mitigation and robust coexistence of multiple FH networks. The main results are summarized as follows.

1. We develop an expectation-maximization (EM) algorithm combined with a 2-D frequency estimator for hop timing and frequency estimation of multiple FH signals with transmitter-receiver bandwidth mismatch. The algorithm resolves frequency collision without retransmission. A software testbed is developed by applying the algorithm in Bluetooth piconets. Simulation results demonstrate the effectiveness of the proposed method.

2. When there is transmitter-receiver bandwidth mismatch, FH signals may hop in and out of the observation band of the receiver, which results in model order variation. We design a low complexity approach for model order variation detection based on the trace of the covariance matrix of the received signals. This approach outperforms the sliding-window singular value decomposition (SVD) or minimum description length (MDL) methods.

3. Multidimensional frequency estimation plays an important role in collision resolution of multiple FH signals. We optimize eigenvector-based multidimensional frequency estimation by minimizing the estimation error variance.

In the following, upper (lower) bold face letters are used for matrices (column vectors). $\boldsymbol{A}^*$, $\boldsymbol{A}^T$, $\boldsymbol{A}^H$, and $\boldsymbol{A}^\dagger$ denote the conjugate, transpose, Hermitian transpose, and pseudo-inverse of $\boldsymbol{A}$, respectively. We use $\otimes$ for the Kronecker product, and $\odot$ for the Khatri-Rao (column-wise Kronecker) product. We also use $\boldsymbol{I}_p$ for a $p \times p$ identity matrix, $\boldsymbol{0}_{M\times N}$ for an $M \times N$ zero matrix, $\boldsymbol{D}(\boldsymbol{a})$ for a diagonal matrix with $\boldsymbol{a}$ as its diagonal, $\boldsymbol{A}^{(m)}$ for a sub-matrix of $\boldsymbol{A}$ formed by its first m rows, and $\|\cdot\|$ for the l_2 norm.

B.4 Collision Resolution Using an Array with Bandwidth Mismatch

Relying on the principle of expectation-maximization and efficient 2-D frequency estimation, we develop an EM algorithm to jointly estimate the hop timing and frequencies of multiple frequency signals without the knowledge of their hop patterns, and it remains operational in the presence of identical frequencies and bandwidth mismatch. The idea behind the collision resolution algorithm is that if a hop-free dataset were available, one could model the collided data packets as a mixture of (modulated) complex exponentials. Cast in proper matrix form, such a signal has a Vandermonde structure in the time domain. In addition, the use of a uniformly spaced linear array, further induces Vandermonde structure in the spatial domain. We exploit the 2-D Vandermonde structures and use a 2-D frequency estimation algorithm that draws upon the rich identifiability and near-optimality results in [28] to recover the hop frequencies. Hop timing estimates are obtained by coupling the 2-D frequency estimation with the expectation maximization algorithm, thus collided packets can be recovered without retransmission.

B.4.1 Data Model

Suppose at time $t \in [t_k, t_{k+1})$, an M-element uniform linear array (ULA) receives a total of d_k effective signals. Each far field FH signal is from a nominal direction-of-arrival (DOA) with negligible angle spread. If there exists transmitter-receiver bandwidth mismatch, the FH signals hop in and out of the observation band of the receiver, which results in model order variations. Here model order variation refers to the change of d_k. The $M \times 1$ baseband received signal vector at the array output can be written as

$$\boldsymbol{x}(t) = \sum_{i=1}^{d_k} \boldsymbol{a}(\theta_{i,k})\beta_{i,k}s_{i,k}(t) + \boldsymbol{v}(t), \quad t_k \le t < t_{k+1}, \tag{1}$$

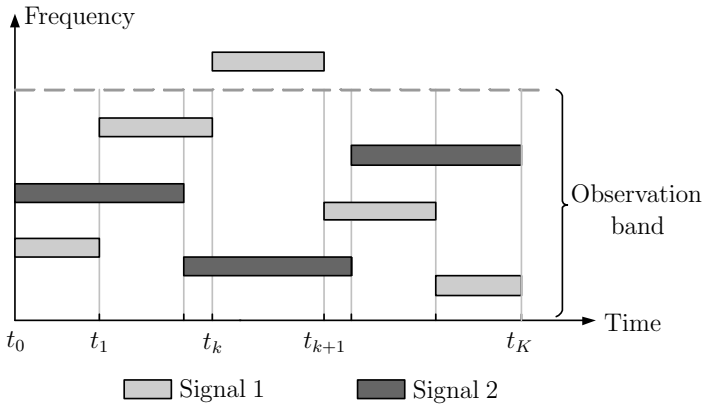


Figure 1: An illustration of two FH signals

where t_k is the k-th system-wide hop instant (hop timing), and d_k is the number of effective signals during the time period delimited in $t_k \leq t < t_{k+1}$. Here $\boldsymbol{a}(\theta_{i,k})$ is the array steering vector, which can be represented as $\boldsymbol{a}(\theta_{i,k}) = \left[1\ \theta_{i,k}\ \cdots\ \theta_{i,k}^{M-1}\right]^T$, and $\theta_{i,k} = e^{j2\pi\Delta\sin(\alpha_{i,k})}$, where $\alpha_{i,k}$ is the DOA of the i-th narrow-band signal during the k-th time period, and Δ is the array baseline separation in terms of wavelength. In (1), $\beta_{i,k}$ denotes the complex path loss for the i-th signal that collects the overall attenuation and propagation phase shift. $\boldsymbol{v}(t)$ is the complex white Gaussian noise vector. For simplicity of exposition, the transmitted FH signal is assumed to be FSK modulated, though the algorithm can also accommodate other modulations such as phase shift keying (PSK) or quadrature amplitude modulation (QAM). The i-th transmitted FH signal is $s_{i,k}(t) = e^{j\omega_{i,k}(t)t}$. The power of $s_{i,k}(t)$ is absorbed into $\beta_{i,k}$. Here the carrier shifts due to hopping or symbol modulation are treated as conceptually equivalent, albeit of different magnitudes.

An illustration of two FH signals is shown in Fig. 1, where t_k denotes a system-wide hop instant. The two FH signals are asynchronous and have different hop rates. The received signal in (1) is hop-free between any two adjacent system-wide hop instants. Signal 1 hops out of the observation band at t_k and hops back in at t_{k+1}, so the number of the effective signals in (1) during $t_k \leq t < t_{k+1}$ is $d_k = 1$, while during other hop-free segments the model order is 2.

After sampling the signals in (1) with a normalized sampling period, the discrete-time equivalent model becomes

$$\boldsymbol{x}_n = \sum_{i=1}^{d_k} \boldsymbol{a}(\theta_{i,k})\beta_{i,k}s_{i,k}(n) + \boldsymbol{v}_n, \quad n_k \leq n < n_{k+1}, \tag{2}$$

where $n = 0, \cdots, N-1$, and N is the total number of snapshots; n_k is the k-th hop instant, where $k = 0, \ldots, K$, and K is the total number of hops. By default we let $n_0 = 0$ and $n_K = N$. The n-th snapshot of the i-th transmitted signal is $s_{i,k}(n) = e^{j\omega_{i,k}\cdot n}$, $n_k \leq n < n_{k+1}$, where $\omega_{i,k}$ is the frequency of the i-th signal during the time period from n_k to $n_{k+1} - 1$. Since n_k is a system-wide hop instant, it is possible that $\omega_{i,k} = \omega_{i,k-1}$ for some i. If $\omega_{i,k} = \omega_{j,k}$, a collision occurs. The objective here is to estimate $\{n_k\}$ and hop frequencies for a given set of observations $\{\boldsymbol{x}_n\}_{n=0}^{N-1}$, with unknown DOAs, hop sequences, amplitudes, and number of effective signals. After these parameters are obtained, collisions are then resolved.

B.4.2 Known Hop Timing

If hop timing is known, i.e., n_k's are known, then the received data set can be split into K hop-free data subsets. Between any two consecutive hop instants, e.g., n_k and n_{k+1}, there are d_k temporal frequencies involved. During such a time segment, the k-th received data subset can be written as

$$\boldsymbol{X}_k = \boldsymbol{A}_k \boldsymbol{B}_k \boldsymbol{S}_k^T + \boldsymbol{V}_k, \tag{3}$$

where $\boldsymbol{X}_k = \left[\, \boldsymbol{x}_{n_k} \, \boldsymbol{x}_{n_k+1} \, \cdots \, \boldsymbol{x}_{n_{k+1}-1} \, \right]$, and $\boldsymbol{A}_k$, $\boldsymbol{B}_k$, $\boldsymbol{S}_k$, and $\boldsymbol{V}_k$ are the corresponding matrices for antenna response, channel attenuation, transmitted signals, and noise. Note that the number of effective signal is d_k (i.e., model order), which can be estimated from $\boldsymbol{X}_k$ by an appropriate source enumeration method because it is fixed in this subset. A source enumeration method can be based on rank detection criteria (e.g. SVD [29]) or information theoretic criteria (e.g., the Akaike information criterion (AIC) [30], MDL [31,32]). SVD requires the knowledge of noise variance for threshold setting in model order selection [29]. AIC tends to overestimate the number of signals in the large sample limit, while MDL is shown to be a consistent estimator [31]. Therefore we choose MDL to estimate d_k from $\boldsymbol{X}_k$, which is to calculate

$$\mathrm{MDL}(m) = -\log \left(\frac{\prod_{i=m+1}^{M} \lambda_i^{\frac{1}{M-m}}}{\frac{1}{M-m} \sum_{i=m+1}^{M} \lambda_i} \right)^{(M-m)(n_{k+1}-n_k)} + \frac{1}{2} m(2M-m) \log(n_{k+1}-n_k), \tag{4}$$

where $\lambda_1 > \lambda_2 \cdots > \lambda_M$ are the eigenvalues of the sample covariance matrix of $\boldsymbol{X}_k$. The number of signals d_k is determined as the value of $m \in \{0, 1, \ldots, M-1\}$ for which $\mathrm{MDL}(m)$ in (4) is minimized. If $d_k > M$, this method does not work. However, exploiting the inherent structure of the data matrix $\boldsymbol{X}_k$, we can use data smoothing to expand the size the $\boldsymbol{X}_k$ while keeping its model order unchanged. Then, the MDL algorithm can be applied to the expanded data matrix to estimate the model order even if $d_k > M$ (see [33] for details).

After d_k is obtained, the estimation of DOAs and frequencies from $\boldsymbol{X}_k$ in (3) can be viewed as a 2-D frequency estimation problem because both $\boldsymbol{A}_k$ and $\boldsymbol{S}_k$ are Vandermonde matrices. There are d_k frequency components along each of the spatial and temporal dimensions of the 2-D frequency mixture $\boldsymbol{X}_k$. Various subspace methods can be used to estimate $\theta_{i,k}$, $\beta_{i,k}$, and $\omega_{i,k}$, for $i = 1, \ldots, d_k$, from $\boldsymbol{X}_k$. For example, eigenvalue-based algorithms such as estimation of signal parameter via rotational invariance technique (ESPRIT) [34], unitary ESPRIT [35], joint angle-frequency estimation (JAFE) [36], matrix enhancement and matrix pencil (MEMP) [37], and eigenvector-based algorithm such as multidimensional folding (MDF) [28], are all applicable. We choose the MDF algorithm here because MDF achieves the most relaxed identifiability bound and is shown to outperform ESPRIT-like algorithms such as JAFE and MEMP, and do not require an extra frequency pairing step (see [28, 38] for more detail). The description of the MDF algorithm can be found in [28, 38] and is omitted here.

B.4.3 Hop Timing Estimation

Both the model order estimation by the MDL algorithm and 2-D frequency estimation by the MDF algorithm need to work with a hop-free data subset. However, here the hop timing is unknown. In

the following, we develop an EM algorithm that yields joint estimates of hop timing and frequencies. It is shown in [39] that the EM algorithm can be written as

E-step: for $1 \leq k \leq K-1$, compute

$$n_k^{(p)} = \arg\min_{n_k} \left\{ \left\| \boldsymbol{X}_k - \boldsymbol{A}_k^{(p)} \boldsymbol{B}_k^{(p)} \left(\boldsymbol{S}_k^{(p)} \right)^T \right\|_F^2 + \left\| \boldsymbol{X}_{k-1} - \boldsymbol{A}_{k-1}^{(p)} \boldsymbol{B}_{k-1}^{(p)} \left(\boldsymbol{S}_{k-1}^{(p)} \right)^T \right\|_F^2 \right\}, \quad (5)$$

where $n_k \in \left[n_{k-1}^{(p)} + 1, n_{k+1}^{(p-1)} - 1 \right]$.

M-step: compute

$$\boldsymbol{\phi}^{(p+1)} = \arg\min_{\boldsymbol{\phi}} \left\{ \sum_{k=0}^{K-1} \left\| \boldsymbol{X}_k - \boldsymbol{A}_k \boldsymbol{B}_k \boldsymbol{S}_k^T \right\|_F^2 \middle| \boldsymbol{n}^{(p)} \right\}. \quad (6)$$

As discussed in Sec. B.4.2, when $\boldsymbol{n}^{(p)}$ is given, the model order can be obtained using the MDL algorithm and other parameters of $\boldsymbol{\phi}^{(p+1)}$ can be obtained by applying the MDF algorithm to the corresponding $\boldsymbol{X}_k$'s determined by $\boldsymbol{n}^{(p)}$. It can be observed from (5) and (6) that the proposed EM approach is actually a sort of decoupled ML algorithm, since $\boldsymbol{n}$ and $\boldsymbol{\phi}$ are estimated in the E-step and M-step respectively in each iteration.

In summary, given a received data block with model order variation and multiple hops, the EM algorithm first takes a randomly generated or pre-determined $\boldsymbol{n}^{(0)}$ as the initialization for the EM algorithm. Then, it iterates the following two steps until convergence: the M-step, Eqn. (6), provides a new estimate of model orders and hop frequencies, which are accomplished by applying first the MDL algorithm and then the MDF algorithm to data segments according to the updated assignment; the E-step, Eqn. (5), involves assigning signal segments to the current estimated parameters that fit them best. Upon convergence, frequencies and complex amplitudes of different segments pertaining to a particular signal can be associated via their corresponding DOA parameters, since for a single data subset, frequency and DOA parameters pertaining to one signal are paired up automatically by the MDF algorithm. Therefore, joint hop timing and frequency estimation for multiple FH signals is achieved and collisions are resolved. Furthermore, to improve the performance, in [39] we also develop a low complexity initialization based on the spectrogram of the received data for the EM algorithm.

B.4.4 Simulation Results

In this section, we first compare the EM algorithm with the iterative ML method of [27], for hop timing estimation in a single-user two-hop setup (one hop timing instant is to be estimated), which is the signal model considered in [27]. Then, we evaluate the performance of the proposed EM algorithm in more general cases.

Performance Comparison for Single-user Two-hop Scenario

For convenience let us denote the iterative ML algorithm of [27] as IML. We compare our EM algorithm to IML. Suppose that the data sequence length is $N = 128$ and the signal of a single

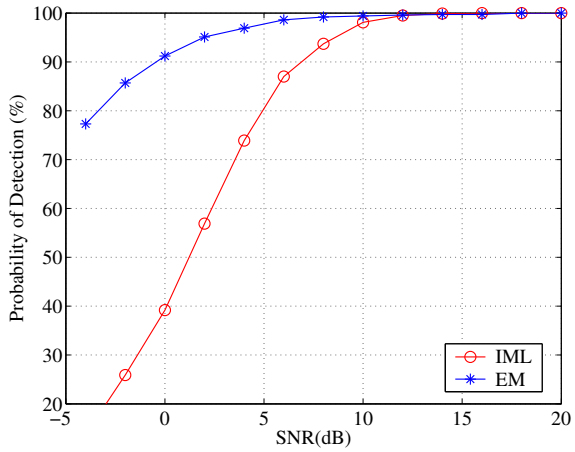


Figure 2: Probability of detection in the single-user scenario.

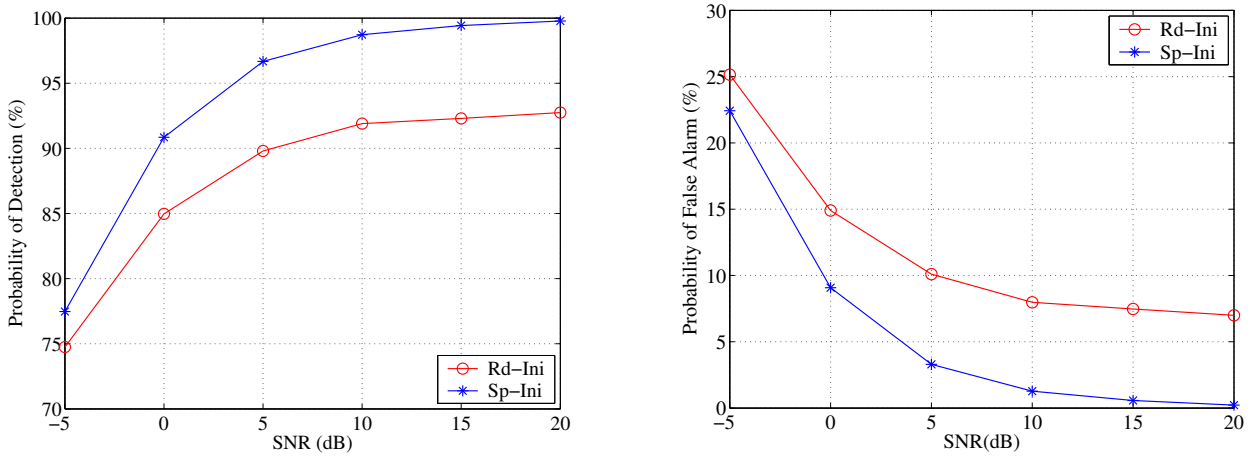
user hops once, from one frequency to another, in this data sequence. The EM algorithm uses a receive array with $M = 6$ antennas, with a baseline separation of half a wavelength of the carrier frequency. IML uses one antenna. Monte Carlo simulations are carried out to compare the average performance of these two algorithms. In each of 1000 realizations, we randomly generate not only the DOA, hop frequencies (ranging from 1 KHz to 58 KHz) and complex amplitudes, but also the hop instant (ranging from the 30-th snapshot to the 98-th snapshot). Note that in [27] the hop instant is fixed around the middle of the data sequence. The probabilities of detection for both algorithms are shown in Fig. 2 (b). Here if the estimated hop instant is the same as the true hop instant, the detection is considered successful. From Fig. 2 (b), we find that the EM algorithm significantly outperforms the IML algorithm at low SNR range. This is due to the following two reasons: (i) the EM algorithm utilizes multiple antennas while IML uses only one antenna; (ii) it is not shown in [27] whether IML can guarantee identifiability for frequency estimation, while the MDF algorithm we used here has the identifiability guarantee.

The complexity order of IML to estimate one hop instant for a single user is $\mathcal{O}(N^2)$ as shown in [27]. The proposed EM algorithm has a complexity order of $\mathcal{O}(KMN^2)$ for multiple signals with K hops. Since $M \ll N$, we conclude that the complexities of the two approaches for one hop instant estimation in single user case are of the same order.

Multiple Signals with Multiple Hops

We evaluate the performance of the EM algorithm for the case of multiple signals with multiple hops. The receiver array has $M = 6$ antennas. The receiver's bandwidth is 60 MHz. Three FH emitters impinge the receiver array with randomly generated DOAs and complex amplitudes. The hop bandwidth of each FH emitter is 80 MHz, which is occupied by 80 distinct frequency channels with 1 MHz channel spacing. Therefore there is a transmitter-receiver bandwidth mismatch. The data size is 6×400. In each realization, hop instants of the three FH signals are randomly generated, and hop frequencies are randomly chosen from the 80 channels.

For the purpose of performance evaluation, a successful detection of hop instant is defined as



(a) (b)

Figure 3: The EM algorithm for hop timing estimation: (a) probability of detection; (b) probability of false alarm.

follows: for a given true hop instant, if there exists an estimated hop instant, whose deviation from the true one is less than or equal to 4 snapshots, then this detection is considered successful. For a given estimated hop instant, if its distances from all true instants are greater than 4 snapshots, a false alarm is claimed to have occurred. The average length of a hop-free segment is 80 snapshots, therefore 4 snapshots represent 5% of that value. Furthermore, the minimum deviation from all estimated hop instants to a given true hop instant is defined as the estimation error for this hop instant. In order to normalize the error, we divide it by the average length of a hop-free segment.

Fig. 3 (a) plots the probability of detection of hop instants using the EM algorithm, where both random ("Rd-ini") and spectrogram-based ("Sp-Ini") initializations are tested. The corresponding probability of false alarm is shown in Fig. 3 (b). The results show that the EM algorithm with spectrogram-based initialization does a good job for hop timing estimation, given the fact that the hop sequences and model order variations are unknown. It can be seen that the probability of detection is about 95% when SNR is at 5 dB, and almost 100% when SNR is greater than 15 dB. Correspondingly, the probability of false alarm is less than 5% at SNR=5 dB, and close to zero when SNR is greater than 15 dB. The EM algorithm with spectrogram-based initialization significantly outperforms its randomly initialized counterpart.

Because of using a simplified EM algorithm, in the p-th iteration, the k-th hop instant n_k is only searched in the time segment bounded by $n_{k-1}^{(p)}$ and $n_{k+1}^{(p-1)}$. Hence it is desirable that the initialization is approximately in-line with each system-wide dwell, which can be achieved by the spectrogram method as long as the separation of two adjacent hop instants is larger than the window size of the spectrogram.

B.4.5 An Application Testbed for Bluetooth

A software testbed is developed to apply the collision resolution method in multiple co-located Bluetooth piconets. The system block diagram is shown in Fig. 4. Transmissions from multiple piconets are received at multiple devices. Without loss of generality, we assume that each piconet

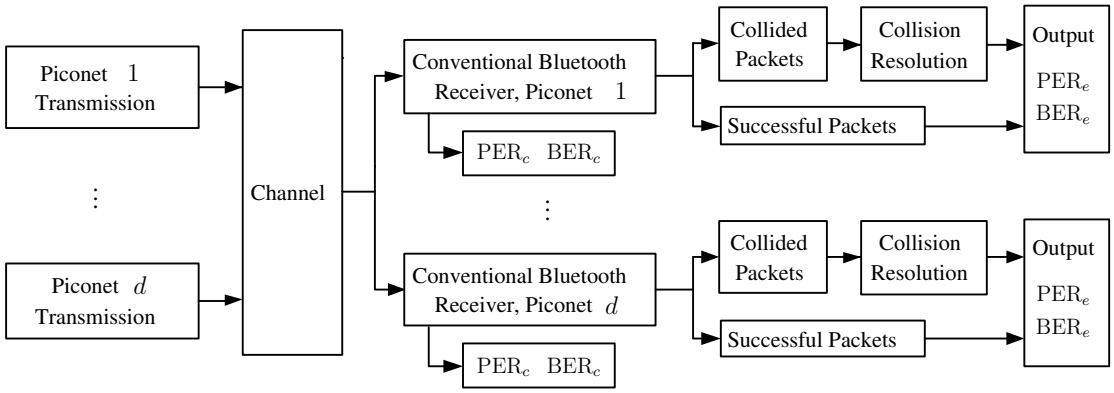


Figure 4: Collision resolution in multiple Bluetooth piconets.

consists of a slave and a master device. As shown in Fig. 4, a user in piconet 1 may receive transmissions from d piconets. The transmissions are not necessarily synchronized across different piconets. The channel is assumed to be flat-fading, though the algorithm can be extended to deal with frequency-selective channels. The proposed receiver consists of a conventional Bluetooth receiver and a collision resolution unit. We assume that packet collision can be detected, and collided data packets are sent to the collision resolution unit when collisions occur, while successful packets are sent directly to the output. For the purpose of performance comparison, we measure the BERs and PERs of both collision-resolution enhanced receiver (BER_e and PER_e), and conventional receiver (BER_c and PER_c). The software testbed for the collision resolution technique is developed based on Matlab Simulink environment with a customized graphic user interface. It simulates data packet transmissions in multiple Bluetooth piconets, and calculates the PER and BER. A number of parameters can be changed in the testbed to simulate difference scenarios, such as the number of piconets, data packet length and SNR.

B.5 Low Complexity Model Order Variation Detection

It is shown in the previous section that when there is transmitter-receiver bandwidth mismatch, model order variation has to be detected before parametric methods can be applied. Relying on a receiver equipped with a uniform linear array (ULA), we have developed a low-complexity method for the detection of model order variations in the context of FH networks with transmitter-receiver bandwidth mismatch [41]. We note that the proposed method is also applicable in other similar scenarios where the transmitted signals are uncorrelated.

For a given set of observations, one may use a sliding-window to obtain data subsets, and subsequently apply source enumeration techniques such as SVD or MDL to each subset to estimate its model order. Model order change points can then be detected by comparing the results as the data window slides. However, there are several disadvantages to this method. First, the complexity of SVD and MDL is high. Second, the resolution of estimated timing is limited to the order of the window size at low SNR, since decisions based on SVD or MDL provide hard-information (the model order). Third, SVD and MDL break down when the number of signals is more than the

number of antennas.

Here we also adopt a sliding-window approach. More importantly, for each data subset, we compute the trace of the signal correlation matrix by exploiting inherent data structure. The trace sequence provide soft-information that later is fed into a recursive or iterative search procedure to obtain accurate timing estimates of model order variations. The proposed method is computationally much simpler than sliding-window SVD or MDL, and it also improves the probability of detection at low SNR. Moreover, the method works even if the number of effective signals is more than the number of sensors.

B.5.1 A Trace Calculation Approach

In this section we illustrate how soft-information is obtained by calculating the trace of the correlation matrix of each data subset. Section B.5.2 discusses the subsequent detection of model order variations based on the soft-information sequence.

The data model is given in (2). Let the window size be P. We assume that the window is small enough so that the model order changes at most once in each window, which is a realistic assumption for slow FH systems, and if the sampling rate is fast enough or the analysis window is short enough, this assumption is also valid for fast FH systems. Without loss of generality, we take a data subset

$$\boldsymbol{X} = [\boldsymbol{x}_q \ \cdots \ \boldsymbol{x}_{q+P-1}],$$

to illustrate the trace calculation approach in the following two cases: constant model order and varying model order within this data subset.

Case 1: Constant Model Order

If the model order is constant and equal to d during the given window, the $M \times P$ data matrix can be written as

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S} + \boldsymbol{W},$$

where $\boldsymbol{A} = [\beta_1 \boldsymbol{a}(\theta_1) \ \beta_2 \boldsymbol{a}(\theta_2) \ \cdots \ \beta_d \boldsymbol{a}(\theta_d)]$, $\boldsymbol{S} = [\boldsymbol{s}_q \ \boldsymbol{s}_{q+1} \ \cdots \ \boldsymbol{s}_{q+P-1}]$, and for $q \le n < q+P$, $\boldsymbol{s}_n = [e^{j\omega_{1,n}\cdot n} \ e^{j\omega_{2,n}\cdot n} \ \cdots \ e^{j\omega_{d,n}\cdot n}]^T$. $\boldsymbol{W}$ is the corresponding noise matrix. Suppose the noise and signal are independent, the correlation matrix of $\boldsymbol{X}$ is given by

$$\boldsymbol{R}_X = E[\boldsymbol{X}\boldsymbol{X}^H] = \boldsymbol{A}\boldsymbol{R}_S\boldsymbol{A}^H + \boldsymbol{R}_W, \tag{7}$$

where $\boldsymbol{R}_S = E[\boldsymbol{S}\boldsymbol{S}^H]$ is the signal correlation matrix, and the noise correlation matrix is $\boldsymbol{R}_W = \sigma_w^2 \boldsymbol{I}_M$. $\boldsymbol{I}_M$ is an $M \times M$ identity matrix. Assuming that signals from different users are uncorrelated, the signal correlation matrix becomes $\boldsymbol{R}_S = \boldsymbol{I}_d$. Hence the trace of $\boldsymbol{R}_X$ is given by

$$\mathrm{tr}(\boldsymbol{R}_X) = \mathrm{tr}(\boldsymbol{A}^H\boldsymbol{A}) + M\sigma_w^2 = M\sum_{i=1}^{d} |\beta_i|^2 + M\sigma_w^2, \tag{8}$$

where we have used the fact that $\boldsymbol{A}$ is a column-scaled Vandermonde matrix. We refer to the trace of $\boldsymbol{R}_X$ as the "soft-information", which is a linear function of the channel power as shown in (8).

Notice that the objective here is to detect the change points, rather than the model order itself. If the change points are known, various source enumeration techniques can be applied, e.g., SVD [29], MDL [31, 32], and PDL [40]. With finite samples, $\widehat{\boldsymbol{R}}_S \approx \boldsymbol{I}_d$, $\widehat{\boldsymbol{R}}_W \approx \sigma_w^2 \boldsymbol{I}_M$, and the correlation matrix of $\boldsymbol{X}$ can be estimated as $\widehat{\boldsymbol{R}}_X = \frac{1}{P}\sum_{n=q}^{q+P-1} \boldsymbol{x}_n \boldsymbol{x}_n^H$.

Case 2: Varying Model Order

We now illustrate the soft-information in the presence of a model order change within the given data subset $\boldsymbol{X}$. For simplicity, assume that a total of d users are active in the FH system, and at time $q+l$, the first user hops out of the observation band. Notice that if more than one users hop out at the same time, the method detects the model order variation more efficiently because of the obviously larger change in (8). We can write the data matrix as $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S} + \boldsymbol{W}$, where

$$\boldsymbol{S} = [\, \boldsymbol{s}_q \ \cdots \ \boldsymbol{s}_{q+l-1} \ \ \boldsymbol{s}_{q+l} \ \cdots \ \boldsymbol{s}_{q+P-1} \,], \tag{9}$$

and for $n < q+l$, $\boldsymbol{s}_n$ is the same as in Case 1, and for $n \geq q+l$, $\boldsymbol{s}_n$ is the same as in Case 1 except that its first element is zero. In this case, $\boldsymbol{R}_s \neq \boldsymbol{I}_d$. In fact, it is shown in [41] that the sample signal correlation matrix is

$$\widehat{\boldsymbol{R}}_S = \frac{1}{P}\sum_{n=q}^{q+P-1} \boldsymbol{s}_n \boldsymbol{s}_n^H \approx \begin{bmatrix} \frac{l}{P} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{d-1} \end{bmatrix}. \tag{10}$$

Similar to (8), we obtain

$$\mathrm{tr}(\boldsymbol{R}_X) = \mathrm{tr}(\boldsymbol{A}\boldsymbol{R}_S\boldsymbol{A}^H) + M\sigma_w^2 = \frac{Ml}{P}|\beta_1|^2 + M\sum_{i=2}^{d}|\beta_i|^2 + M\sigma_w^2. \tag{11}$$

It can be seen that the trace of the estimated correlation matrix with base at $n = q$ is linear in l, the delay or lag to the change point. If there are multiple changes within the window, the situation is more complicated. For example, an incoming signal and a vanishing signal in the same window could cancel each other's effect on the model order. Therefore we assume that the window size is small enough that the model order changes at most once within each window.

The previous derivation assumes that $\boldsymbol{A}$ is a Vandermonde matrix. If $\boldsymbol{A}$ is not a Vandermonde matrix, then in case of fixed model order, (8) becomes $\mathrm{tr}(\boldsymbol{R}_X) = \sum_{i=1}^{d} ||\boldsymbol{a}_i||_2^2 + M\sigma_w^2$, where $\boldsymbol{a}_i$ is the i-th column of $\boldsymbol{A}$, and $||\cdot||_2$ stands for vector 2-norm. In the case of varying model order, (11) becomes $\mathrm{tr}(\boldsymbol{R}_X) = \frac{l}{P}||\boldsymbol{a}_1||_2^2 + \sum_{i=2}^{d}||\boldsymbol{a}_i||_2^2 + M\sigma_w^2$. Hence the trace is also linear to the "delay" to the change point, and the proposed method is still applicable.

In summary, given the observations $\boldsymbol{x}(n)$, $0 \leq n < N$, we can obtain a soft-information sequence by using a sliding-window as follows $y_q = \mathrm{tr}\left(\sum_{p=q}^{q+P-1} \boldsymbol{x}_p \boldsymbol{x}_p^H\right)$, $q = 0, \ldots, Q$, where P is the window size and $Q = N - P$. Let us use an example to illustrate $\boldsymbol{y} = [y_0, \ldots, y_Q]^T$. Suppose a data set with 600 snapshots is available, and the number of effective signals in the observation band first changes from three to one, then from one to two, due to frequency hopping. The change points are 170 and 410. The sequence $\boldsymbol{y}$ is shown in Fig. 5, where the window size is 50 and

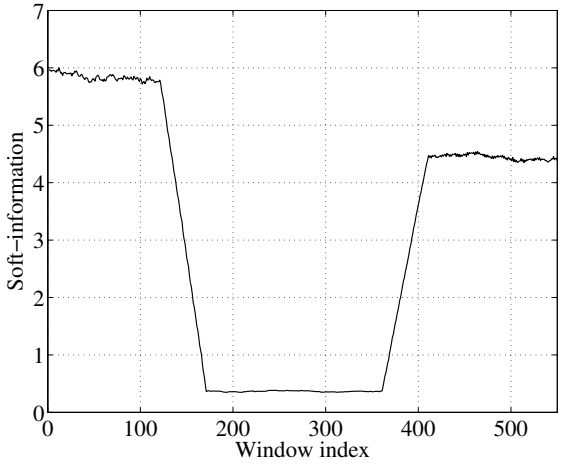


Figure 5: Example of the soft-information sequence obtained by trace calculation using a sliding-window, where true points of model order changes are 170 and 410.

SNR is 10 dB. SNR is defined as $10\log_{10}(1/\sigma_w^2)$. In Fig. 5, we observe that a model order change manifests itself as a linear variation of the soft-information. The change instants are given by the points at which the soft-information flattens out.

Comparing to source enumeration algorithms such as SVD and MDL, the trace calculation method is of much lower complexity, and works even when the number of the signals is more than the number of antennas. For example, for an $M \times P$ data subset, the complexity is about $\mathcal{O}(MP)$ for trace calculation, and $\mathcal{O}(M^2P + P^3)$ for SVD and MDL. After obtaining the soft-information sequence $\boldsymbol{y}$, the next step is to estimate the change points, n_k, for $k = 0, \cdots, K-1$. This is discussed next.

B.5.2 Change Detection Methods

We have developed three methods to estimate the change points from a sequence $\boldsymbol{y}$, which can be a soft-information sequence obtained by the trace calculation or a hard-information sequence obtained by SVD or MDL. One method is a recursive search similar to the dynamic programming principle (see, e.g., [42, Chapter 12]), and the other is a decoupled iterative search similar to the expectation maximization principle (see, e.g., [43]). However, we note that since the detection is not maximum likelihood, and we do not pursue dynamic programming nor expectation maximization algorithms. The third method is a simple low-complexity difference approach to estimate the change points from the soft-information sequence. More details can be found in [41].

B.5.3 Simulation Results

We present simulation results to illustrate the performance of the proposed methods. For a given data matrix with model order variations, we apply the trace calculation (TC), SVD, and MDL approaches to data subsets obtained using a sliding-window. Subsequently the recursive, the iterative and the difference detection methods are used to estimate the change points. Notice that SVD and

Table 1: Probability of detection of model order variation (AWGN channels)

SNR		0 dB	3 dB	6 dB	9 dB
P_d %	SVD-R	3	28	89	98
	SVD-I	3	28	87	98
	MDL-R	7	42	91	99
	MDL-I	6	42	90	99
	TC-R	51	71	91	97
	TC-I	36	53	81	94
	TC-D	82	91	97	98

Table 2: Probability of detection of model order variation (flat-fading channels)

SNR		0 dB	3 dB	6 dB	9 dB
P_d %	SVD-R	33	51	83	95
	SVD-I	30	50	84	94
	MDL-R	35	55	87	95
	MDL-I	33	55	86	95
	TC-R	59	72	84	91
	TC-I	46	59	75	85
	TC-D	80	87	92	94

MDL return hard-information – the actual decision on model orders for each data subset. The decision threshold used in the SVD method is chosen according to the modified third bound given in [29]. The MDL algorithm used here follows [31].

Monte Carlo simulations are conducted to assess the performance of the proposed method. The sources are three narrow-band FH emitters with randomly generated DOAs. We generate 200 realizations and compute the probability of detection for each SNR. In each realization, the model order change points and the number of signals that hop in/out of the observation band are randomly generated. A detection is defined successful when the difference between an estimated instant and the true one is less than or equal to 5, which is ten percent of the window length 50. Table 1 shows the probability of detection (P_d) using the recursive method (R), the iterative method (I) and the difference method (D) for AWGN channels, where all $\beta_{i,k}$'s are fixed to 1. Table 2 shows the probability of detection using these three methods for flat-fading channels, where all $\beta_{i,k}$'s are randomly generated with standard complex Gaussian distribution.

For both recursive and iterative detections, the results indicate that the TC based approach outperforms the SVD and MDL based methods at low SNR regime. This is mainly because that at low SNR, the resolution of SVD and MDL is limited by the window size, while TC does not

have this problem since it returns soft-information. The iterative detector slightly underperforms the recursive detector but comes with much lower complexity. The difference method combined with trace calculation has the least complexity and competitive performance. When SNR is greater than 10 dB, all three methods (SVD, MDL, and TC) can achieve a probability of detection close or equal to 100%.

B.6 Optimizing Eigenvector-Based Frequency Estimation

Multidimensional frequency estimation plays an important role in our proposed interference mitigation framework. Recently an eigenvector-based algorithm has been developed for multidimensional frequency estimation with a single snapshot of data mixture. Unlike most existing algebraic approaches that estimate frequencies from eigenvalues, the eigenvector-based algorithm achieves automatic frequency pairing without joint diagonalization of multiple matrices, but it fails when there exist identical frequencies in certain dimensions because eigenvectors are not linearly independent anymore. We develop an eigenvector-based algorithm for multidimensional frequency estimation with finite data snapshots. We introduce complex weighting factors so that the algorithm is still operational when there exist identical frequencies in one or more dimensions. Furthermore, the weighting factors are optimized to minimize the mean square errors of the frequency estimates. Simulation results show that the proposed algorithm offers competitive performance when compared to existing algebraic approaches but at lower complexity.

The data model for N-D frequency estimation can take a variety of forms. Here we refer to a single snapshot N-D frequency mixture as an N-D array $\underline{\boldsymbol{X}}$ with typical element

$$x_{m_1,m_2,\cdots,m_N} = \sum_{f=1}^{F} c_f \prod_{n=1}^{N} e^{j\omega_{f,n}(m_n-1)} + w_{m_1,m_2,\cdots,m_N}, \tag{12}$$

where $m_n = 1,\ldots,M_n$, for $n = 1,\ldots,N$, and M_n is the dimension size of the n-th dimension. The total sample size is $M := \prod_{n=1}^{N} M_n$. In (12), the frequencies $\omega_{f,n} \in (-\pi,\pi]$, for $f = 1,\ldots,F$, $n = 1,\ldots,N$, and $w_{m_1,m_2,\cdots,m_N}$ is white Gaussian noise with variance σ^2. Similarly T snapshots of N-D frequency data mixtures may be modeled as T N-D arrays, $\underline{\boldsymbol{X}}(t)$, with typical element

$$x_{m_1,m_2,\cdots,m_N}(t) = \sum_{f=1}^{F} c_f(t) \prod_{n=1}^{N} e^{j\omega_{f,n}(m_n-1)} + w_{m_1,m_2,\cdots,m_N}(t), \quad t = 1,\cdots,T, \tag{13}$$

where t is the snapshot index, which can be a time index, or trial index in case of multiple trials of experiments. $T = 1$ corresponds to the single snapshot case. The frequency set $\{\omega_{f,n}\}_{n=1}^{N}$ is a N-D frequency component, and there are F such components. The objective of N-D frequency estimation is to estimate $\{\omega_{f,n}\}_{n=1}^{N}$, for $f = 1,\ldots,F$, from given $\underline{\boldsymbol{X}}(t)$, $t = 1,\cdots,T$. Notice that the same data model for multiple snapshot case has been used in [44, 45].

In order to facilitate the presentation, we introduce the equivalent data models based on Khatri-Rao product. Given (13), define the sample vector $\boldsymbol{x}(t)$ for the t-th snapshot as

$$\boldsymbol{x}(t) = \left[x_{1,1,\cdots,1}(t)\; x_{1,1,\ldots,2}(t)\; \cdots\; x_{1,1,\ldots,M_N}(t)\; x_{1,1,\ldots,2,1}(t)\; \cdots\; x_{M_1,M_2,\ldots,M_N}(t)\right]^T.$$

Furthermore, define N Vandermonde matrices $\boldsymbol{A}_n \in \mathbb{C}^{M_n \times F}$ with generators $\{e^{j\omega_{f,n}}\}_{f=1}^{F}$ such that

$$\boldsymbol{A}_n := [\boldsymbol{a}_{1,n}\, \boldsymbol{a}_{2,n}\, \cdots\, \boldsymbol{a}_{F,n}], \text{ where } \boldsymbol{a}_{f,n} = \left[1\, e^{j\omega_{f,n}}\, \cdots\, e^{j(M_n-1)\omega_{f,n}}\right]^T, \quad n = 1, \ldots, N.$$

It can be verified that

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{c}(t) + \boldsymbol{w}(t), \quad t = 1, \ldots, T, \tag{14}$$

where $\boldsymbol{w}(t)$ is the noise vector, and

$$\boldsymbol{A} := \boldsymbol{A}_1 \odot \boldsymbol{A}_2 \odot \cdots \odot \boldsymbol{A}_N, \qquad \boldsymbol{c}(t) := \left[c_1(t)\, c_2(t)\, \cdots\, c_F(t)\right]^T.$$

Define

$$\boldsymbol{X} := \left[\boldsymbol{x}(1)\, \boldsymbol{x}(2)\, \cdots\, \boldsymbol{x}(T)\right] \in \mathbb{C}^{M \times T}, \qquad \boldsymbol{C} := \left[\boldsymbol{c}(1)\, \boldsymbol{c}(2)\, \cdots\, \boldsymbol{c}(T)\right] \in \mathbb{C}^{F \times T},$$

then the data model in (14) can be rewritten in matrix form as

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{C} + \boldsymbol{W}, \tag{15}$$

where $\boldsymbol{W}$ is the corresponding noise matrix. We will need the following lemmas [46].

Lemma 1 *Define a set of N-D selection matrices as*

$$\boldsymbol{J}_{\ell_1,\ell_2,\cdots,\ell_N} := \boldsymbol{J}_{\ell_1}^{K_1} \otimes \boldsymbol{J}_{\ell_2}^{K_2} \cdots \otimes \boldsymbol{J}_{\ell_N}^{K_N}, \tag{16}$$

$$\boldsymbol{J}_{\ell_n}^{K_n} := \left[\boldsymbol{0}_{K_n \times (\ell_n - 1)}\ \boldsymbol{I}_{K_n}\ \boldsymbol{0}_{K_n \times (L_n - \ell_n)}\right], \tag{17}$$

where $\ell_n = 1, \ldots, L_n$, and K_n and L_n are positive integers satisfying $K_n + L_n = M_n + 1$ for $n = 1, \ldots, N$. Further define an N-D smoothing operator for the snapshot vector in (14) *as*

$$\mathcal{S}[\boldsymbol{x}(t)] := \left[\boldsymbol{J}_{1,1,\cdots,1}\boldsymbol{x}(t)\, \boldsymbol{J}_{1,1,\cdots,2}\boldsymbol{x}(t)\, \cdots\, \boldsymbol{J}_{1,1,\cdots,L_N}\boldsymbol{x}(t)\, \boldsymbol{J}_{1,1,\cdots,2,1}\boldsymbol{x}(t)\, \cdots\, \boldsymbol{J}_{L_1,L_2,\cdots,L_N}\boldsymbol{x}(t)\right],$$

then it can be verified that in the absence of noise

$$\boldsymbol{X}_{\mathcal{S}}(t) := \mathcal{S}[\boldsymbol{x}(t)] \ = \left(\boldsymbol{A}_1^{(K_1)} \odot \boldsymbol{A}_2^{(K_2)} \odot \cdots \odot \boldsymbol{A}_N^{(K_N)}\right) \boldsymbol{D}\big(\boldsymbol{c}(t)\big) \left(\boldsymbol{A}_1^{(L_1)} \odot \boldsymbol{A}_2^{(L_2)} \odot \cdots \odot \boldsymbol{A}_N^{(L_N)}\right)^T.$$

Lemma 2 *Given N Vandermonde matrices $\boldsymbol{A}_n \in \mathbb{C}^{M_n \times F}$, with generators $\{e^{j\omega_{f,n}}\}_{f=1}^{F}$, for $n = 1, \ldots, N$, and a complex matrix $\boldsymbol{C} \in \mathbb{C}^{F \times T}$, if we define*

$$\boldsymbol{B} := \begin{bmatrix} \boldsymbol{C}^T \\ \boldsymbol{\Pi}_T \boldsymbol{C}^H \boldsymbol{D}(\boldsymbol{\beta}) \end{bmatrix}, \tag{18}$$

where $\boldsymbol{\Pi}_T$ is a $T \times T$ permutation matrix with ones on its anti-diagonal, and

$$\boldsymbol{\beta} = \left[e^{-j\beta_1}\ \ e^{-j\beta_2}\ \ \cdots\ \ e^{-j\beta_F}\right]^T,$$

$$\beta_f = \sum_{n=1}^{N} (M_n - 1)\omega_{f,n}, \tag{19}$$

then the rank of the matrix

$$\widetilde{\boldsymbol{H}} := \boldsymbol{A}_1^{(L_1)} \odot \boldsymbol{A}_2^{(L_2)} \cdots \odot \boldsymbol{A}_N^{(L_N)} \odot \boldsymbol{B} \tag{20}$$

is $\min\left\{2T\prod_{n=1}^{N} L_n, F\right\}$ almost surely, provided that the NF frequencies $\{\omega_{f,n}\}_{n=1}^{N}$, $f = 1, \cdots, F$, and the TF elements of $\boldsymbol{C}$, are drawn from distributions that are continuous with respect to the Lebesgue measure in $\emptyset^{NF \times 1}$ and $\mathbb{C}^{TF \times 1}$, respectively.

B.6.1 The Eigenvector-Based Algorithm for N-D Frequency Estimation

In this section we present the algorithm for N-D frequency estimation from multiple snapshots. For simplicity of exposition, the algorithm is developed in the noiseless case. The noise effect on the performance of the algorithm is analyzed in Section B.6.2.

Given (14) in the noiseless case, we can apply the smooth operator $\mathcal{S}$ defined in Lemma 1 to every snapshot $\boldsymbol{x}(t)$, and obtain

$$\boldsymbol{X}_{\mathcal{S}}(t) := \mathcal{S}[\boldsymbol{x}(t)] = \boldsymbol{G}\boldsymbol{D}\big(\boldsymbol{c}(t)\big)\boldsymbol{H}^T, \tag{21}$$

where

$$\boldsymbol{G} := \boldsymbol{A}_1^{(K_1)} \odot \boldsymbol{A}_2^{(K_2)} \cdots \odot \boldsymbol{A}_N^{(K_N)}, \quad \boldsymbol{H} := \boldsymbol{A}_1^{(L_1)} \odot \boldsymbol{A}_2^{(L_2)} \cdots \odot \boldsymbol{A}_N^{(L_N)}.$$

The positive integers K_n and L_n, $n = 1, \cdots, N$, are chosen such that

$$K_n + L_n = M_n + 1, \quad 1 \le n \le N. \tag{22}$$

To further explore the data structure, we can perform the forward-backward smoothing on the data vector $\boldsymbol{x}(t)$ in (14). Define $\boldsymbol{y}(t) := \boldsymbol{\Pi}_M \boldsymbol{x}^*(t)$, where $\boldsymbol{\Pi}_M$ is an $M \times M$ permutation matrix with ones on its anti-diagonal. It can be verified that $\boldsymbol{y}(t) = \boldsymbol{A}\widetilde{\boldsymbol{c}}(t)$, where $\widetilde{\boldsymbol{c}}(t) = [\widetilde{c}_1(t), \widetilde{c}_2(t), \cdots, \widetilde{c}_F(t)]^T$, with $\widetilde{c}_f(t) = c_f^*(t)e^{-j\beta_f}$, and β_f is defined in (19). Applying the same technique to $\boldsymbol{y}(t)$ that we used to construct $\boldsymbol{X}_{\mathcal{S}}(t)$ from $\boldsymbol{x}(t)$, we obtain

$$\boldsymbol{Y}_{\mathcal{S}}(t) := \mathcal{S}[\boldsymbol{y}(t)] = \boldsymbol{G}\boldsymbol{D}\big(\widetilde{\boldsymbol{c}}(t)\big)\boldsymbol{H}^T.$$

We then collect all the smoothed data matrices to obtain

$$\widetilde{\boldsymbol{X}} := \begin{bmatrix} \boldsymbol{X}_{\mathcal{S}}(1) & \boldsymbol{X}_{\mathcal{S}}(2) & \cdots & \boldsymbol{X}_{\mathcal{S}}(T) & \boldsymbol{Y}_{\mathcal{S}}(T) & \boldsymbol{Y}_{\mathcal{S}}(T-1) & \cdots & \boldsymbol{Y}_{\mathcal{S}}(1) \end{bmatrix}. \tag{23}$$

It can be verified that

$$\widetilde{\boldsymbol{X}} = \boldsymbol{G}(\boldsymbol{H} \odot \boldsymbol{B})^T = \boldsymbol{G}\widetilde{\boldsymbol{H}}^T, \tag{24}$$

where $\boldsymbol{B}$ and $\widetilde{\boldsymbol{H}}$ are defined in (18) and (20), respectively. A key step of our algorithm is the construction of $\widetilde{\boldsymbol{X}}$ to ensure that it is of rank F almost surely. Note that similar smoothing technique has been used in [44], but its relation to the Khatri-Rao product is not explored. In (24), since $\boldsymbol{G}$ is the Khatri-Rao product of multiple Vandermonde matrices, $\boldsymbol{G}$ is almost surely full column rank if $\prod_{n=1}^N K_n \ge F$. According to Lemma 2, if $2T\prod_{n=1}^N L_n \ge F$, $\widetilde{\boldsymbol{H}}$ has full column rank almost surely. According to the Sylvester's inequality [47]

$$\operatorname{rank}(\boldsymbol{G}) + \operatorname{rank}\big(\widetilde{\boldsymbol{H}}^T\big) - F \le \operatorname{rank}\big(\boldsymbol{G}\widetilde{\boldsymbol{H}}^T\big) \le \min\{\operatorname{rank}(\boldsymbol{G}), \operatorname{rank}\big(\widetilde{\boldsymbol{H}}^T\big)\},$$

hence $\widetilde{\boldsymbol{X}}$ is of rank F almost surely. The singular value decomposition (SVD) of $\widetilde{\boldsymbol{X}}$ yields

$$\widetilde{\boldsymbol{X}} = \boldsymbol{U}_s\boldsymbol{\Sigma}_s\boldsymbol{V}_s^H, \tag{25}$$

where $\boldsymbol{U}_s$ has F columns that together span the column space of $\widetilde{\boldsymbol{X}}$. Since the same space is spanned by the columns of $\boldsymbol{G}$, there exists an $F \times F$ nonsingular matrix $\boldsymbol{T}^{-1}$ such that

$$\boldsymbol{U}_s = \boldsymbol{G}\boldsymbol{T}^{-1}. \tag{26}$$

Similar to the IMDF algorithm [48], once the signal subspace $\boldsymbol{U}_s$ is obtained, we can construct two matrices from $\boldsymbol{U}_s$ whose general eigenvalues are the exponentials of the first dimension, and then use the general eigenvectors to estimate N-D frequencies. However, as mentioned before, the IMDF algorithm fails when there exist identical frequencies in the first dimension since the eigenvectors are not linearly independent anymore. Furthermore, it has been shown in [48] that the performance of the IMDF algorithm is severely degraded if there are close frequencies in the first dimension. To address this problem, in the following, we present a method to construct two matrices whose general eigenvalues are weighted sum of the N-D exponentials. The N-D frequencies are still resolved from the general eigenvectors.

We define two selection matrices $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ as

$$\boldsymbol{J}_1 := \boldsymbol{J}_{1,1} \otimes \boldsymbol{J}_{1,2} \cdots \otimes \boldsymbol{J}_{1,N}, \quad \boldsymbol{J}_2 := \sum_{n=1}^{N} \alpha_n \widetilde{\boldsymbol{J}}_{2,n}, \tag{27}$$

where $\boldsymbol{J}_{1,n} = \left[\boldsymbol{I}_{K_n-1}\ \boldsymbol{0}_{(K_n-1)\times 1}\right]$, $\boldsymbol{J}_{2,n} = \left[\boldsymbol{0}_{(K_n-1)\times 1}\ \boldsymbol{I}_{K_n-1}\right]$, $\widetilde{\boldsymbol{J}}_{2,n} = (\boldsymbol{J}_{1,1} \otimes \cdots \otimes \boldsymbol{J}_{1,n-1}) \otimes \boldsymbol{J}_{2,n} \otimes (\boldsymbol{J}_{1,n+1} \otimes \cdots \otimes \boldsymbol{J}_{1,N})$.

Here $\{\alpha_n\}_{n=1}^N$ are complex weighting factors, which can be randomly chosen initially. As we will show in Section B.6.2, the MSEs of the frequency estimates are affected by these weighting factors in the noisy case. Next, we obtain two equal-sized matrices $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ by

$$\boldsymbol{U}_1 := \boldsymbol{J}_1\boldsymbol{U}_s, \quad \boldsymbol{U}_2 := \boldsymbol{J}_2\boldsymbol{U}_s. \tag{28}$$

According to the property of Khatri-Rao product [49]: $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \odot \boldsymbol{D}) = \boldsymbol{A}\boldsymbol{C} \odot \boldsymbol{B}\boldsymbol{D}$, it can be verified that

$$\boldsymbol{U}_1 = \boldsymbol{P}\boldsymbol{T}^{-1}, \quad \boldsymbol{U}_2 = \boldsymbol{P}\boldsymbol{D}(\boldsymbol{\zeta})\boldsymbol{T}^{-1}. \tag{29}$$

where $\boldsymbol{P} = \boldsymbol{A}_1^{(K_1-1)} \odot \boldsymbol{A}_2^{(K_2-1)} \odot \cdots \odot \boldsymbol{A}_N^{(K_N-1)}$. It is clear that $\boldsymbol{P}$ has full column rank almost surely if $F \leq \prod_{n=1}^N (K_n - 1)$. In (29), $\boldsymbol{\zeta} := [\zeta_1, \zeta_2, \cdots, \zeta_F]^T$, and $\zeta_f = \sum_{n=1}^N \alpha_n e^{j\omega_{f,n}}$. Therefore we have

$$\boldsymbol{U}_1^\dagger \boldsymbol{U}_2 = \boldsymbol{T}\boldsymbol{D}(\boldsymbol{\zeta})\boldsymbol{T}^{-1}.$$

Clearly we can choose $\{\alpha_n\}_{n=1}^N$ to ensure the elements of $\boldsymbol{\zeta}$ are distinct even if there exist identical frequencies in one or more dimensions, but this is not guaranteed by randomly generated $\{\alpha_n\}_{n=1}^N$. We will discuss how to choose the weighting factors in Section B.6.2. $\boldsymbol{T}$ can be retrieved from the eigen-decomposition of $\boldsymbol{U}_1^\dagger \boldsymbol{U}_2$ up to column permutation and scaling ambiguity. Suppose that the eigen-decomposition of $\boldsymbol{U}_1^\dagger \boldsymbol{U}_2$ gives $\boldsymbol{T}_{\text{sp}} = \boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{\Delta}$, where $\boldsymbol{\Lambda}$ is a nonsingular diagonal column scaling matrix and $\boldsymbol{\Delta}$ is a permutation matrix. Once we obtain $\boldsymbol{T}_{\text{sp}}$, we can retrieve $\boldsymbol{P}$ up to column permutation and scaling ambiguity according to

$$\boldsymbol{P}_{\text{sp}} = \boldsymbol{U}_1\boldsymbol{T}_{\text{sp}} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{\Delta}. \tag{30}$$

Notice that $\boldsymbol{P}$ is the Khatri-Rao product of N Vandermonde matrices, and there are F columns in $\boldsymbol{P}$. The N frequencies of the same N-D component appear in the same column of $\boldsymbol{P}$. In other words, for fixed f, $\{\omega_{f,n}\}_{n=1}^{N}$ appear in the same column of $\boldsymbol{P}$. Thanks to this structure, we can obtain F N-D frequency components by dividing suitably chosen elements of the aforementioned columns of $\boldsymbol{P}_{\rm sp}$. Therefore the column scaling and permutation will not have a material effect on the algorithm. For this reason, we may drop subscript "$_{\rm sp}$" from now on as long as it is clear from the context. Suppose $\{e^{j\omega_{f,n}}\}_{n=1}^{N}$ appear in the f-th column of $\boldsymbol{P}$, then each of them can be obtained by anyone of the following quotients

$$e^{j\omega_{f,n}} = \frac{p_{k,f}}{p_{k-K'_n,f}}, \qquad \mod (k-1, K'_{n-1}) \geq K'_n, \text{ for } f = 1, \ldots, F, \tag{31}$$

where $1 \leq k \leq K'_0$, $p_{k,f}$ is the (k,f)-th element of $\boldsymbol{P}$, and

$$K'_n := \begin{cases} \prod_{p=n+1}^{N}(K_p - 1), & 0 \leq n \leq N-1, \\ 1, & n = N. \end{cases} \tag{32}$$

Notice that the frequencies are automatically paired because the frequencies $(\omega_{f,n}, n = 1, \ldots, N)$ of the same N-D component (the f-th component) are obtained from the same column of $\boldsymbol{P}$.

If the data observations are noisy as given in (14), applying the above algorithm we can obtain the estimate of $\boldsymbol{P}$ as $\widehat{\boldsymbol{P}}$. In order to reduce the MSEs of frequency estimates, we use the average of all the quotients in (31) to obtain an estimate of the N-D exponential. Therefore, $e^{j\omega_{f,n}}$ can be estimated by the following average

$$\widehat{e^{j\omega_{f,n}}} = \frac{1}{\mu_n} \sum_{\substack{k=1 \\ \mod\left(k-1, K'_{n-1}\right) \geq K'_n}}^{K'_0} \frac{\hat{p}_{k,f}}{\hat{p}_{k-K'_n,f}}, \quad n = 1, \ldots, N, \tag{33}$$

where $\mu_n = K'_0(K_n - 2)/(K_n - 1)$. The average is also the so-called "circular mean" in direction statistics [50]. Finally the frequency estimates are obtained by

$$\widehat{\omega}_{f,n} = \mathcal{I}\left(\log \widehat{e^{j\omega_{f,n}}}\right). \tag{34}$$

After the frequency estimates are obtained, the amplitude matrix $\boldsymbol{C}$ can be obtained by solving (15) using a least-squares approach.

For an N-D frequency estimation algorithm, the maximum number of uniquely resolvable frequency components in the absence of noise is referred to as its identifiability bound. We summarize our main result on statistical identifiability for the proposed algorithm in the following theorem [46].

Theorem 1 *Given T snapshots of sums of F N-D exponentials as in* (13)*, in the absence of noise, the parameter set $(\{\omega_{f,n}\}_{n=1}^{N}, \{c_f(t)\}_{t=1}^{T})$, $f = 1, \ldots, F$, is almost surely uniquely identifiable by the proposed algorithm in Section B.6.1, provided that*

$$F \leq \max_{\substack{K_n + L_n = M_n + 1 \\ 1 \leq K_n \leq M_n \\ 1 \leq n \leq N}} \min\left(\prod_{n=1}^{N}(K_n - 1),\ 2T\prod_{n=1}^{N} L_n\right), \tag{35}$$

where the NF frequencies $\{\omega_{f,n}\}_{n=1}^{N}$, $f = 1, \ldots, F$, and TF complex amplitudes $\{c_f(t)\}_{t=1}^{T}$, $f = 1, \ldots, F$, are assumed to be drawn from distributions that are continuous with respect to the Lebesgue measure in $\emptyset^{NF\times 1}$ and $\mathbb{C}^{TF\times 1}$, respectively.

B.6.2 Optimization of the Eigenvector-Based Algorithm

We have derived the theoretic error variance of the eigenvector-based algorithm in [46]. We have also derived the Cramér-Rao Bound for multidimensional frequency estimation models [46]. Suppose that $\widehat{\omega}_{f,n} = \omega_{f,n} + \Delta\omega_{f,n}$. It is shown in [46] that

$$\eta = \sum_{n=1}^{N}\sum_{f=1}^{F} \lim_{\sigma^2\to 0} \frac{E[\Delta\omega_{f,n}^2]}{\mathrm{var}_{\mathrm{CRB}}(\widehat{\omega}_{f,n})} \quad \propto \quad \gamma(\boldsymbol{\alpha}), \tag{36}$$

where $\gamma(\boldsymbol{\alpha}) = \sum_{f=1}^{F-1}\sum_{g=f+1}^{F} \frac{1}{|(\boldsymbol{\omega}_f^T - \boldsymbol{\omega}_g^T)\boldsymbol{\alpha}|}$, and $\boldsymbol{\omega}_f = [\omega_{f,1}\ \omega_{f,2}\ \cdots \omega_{f,N}]^T$. To minimize the error variance, an optimal $\boldsymbol{\alpha}$ can be obtained by

$$\boldsymbol{\alpha}_{\mathrm{opt}} = \arg\min_{\boldsymbol{\alpha}} \gamma(\boldsymbol{\alpha}), \ \text{ subject to } \ \|\boldsymbol{\alpha}\| \leq 1. \tag{37}$$

The optimization problem (37) is a so called sum-of-ratios fractional programming problem, which is a difficult global optimization problem [51]. There is no efficient algorithm available to solve it to date. We propose to use grid search in the super-sphere $\|\boldsymbol{\alpha}\| \leq 1$ to find a moderate initial value of $\boldsymbol{\alpha}$, then use a Newton type algorithm to find an optimal $\{\alpha_n\}_{n=1}^{N}$. In order to reduce the complexity, we may set $|\alpha_n| = \sqrt{\frac{1}{N}}$, for $n = 1, \cdots, N$, and the search grid does not need to be fine (for example, the step size of angle in one dimension can be set to π/F).

Alternatively we can use the following method to get a moderate initial value of $\{\alpha_n\}_{n=1}^{N}$. If we define

$$\epsilon(\boldsymbol{\alpha}) := \min_{1\leq f<g\leq F} \left\|(\boldsymbol{\omega}_f^T - \boldsymbol{\omega}_g^T)\boldsymbol{\alpha}\right\|, \tag{38}$$

then we have

$$\gamma(\boldsymbol{\alpha}) \leq \frac{F(F-1)}{\epsilon(\boldsymbol{\alpha})}.$$

If we can solve the following optimization problem

$$\boldsymbol{\alpha}_0 = \arg\max_{\|\boldsymbol{\alpha}\|\leq 1} \epsilon(\boldsymbol{\alpha}), \tag{39}$$

the upper bound of $\gamma(\boldsymbol{\alpha})$ is minimized. The optimization problem (39) can be solved using a sequential quadratic programming (SQP) method, which is a common quasi-Newton type algorithm available in many optimization packages such as the optimization toolbox in Matlab.

The algorithm using an optimal $\boldsymbol{\alpha}$ for N-D frequency estimation is described in Table 3. Notice that the first step, which involves the SVD of $\widetilde{\boldsymbol{X}}$, is most computationally complex. But this step is executed only once. To appreciate the proposed algorithm in Table 3, we compare three methods to obtain an optimal $\boldsymbol{\alpha}$. The difference is only in Step 3 of Table 3, where we may: (a)

Table 3: An improved eigenvector-based algorithm using optimal weighting factors

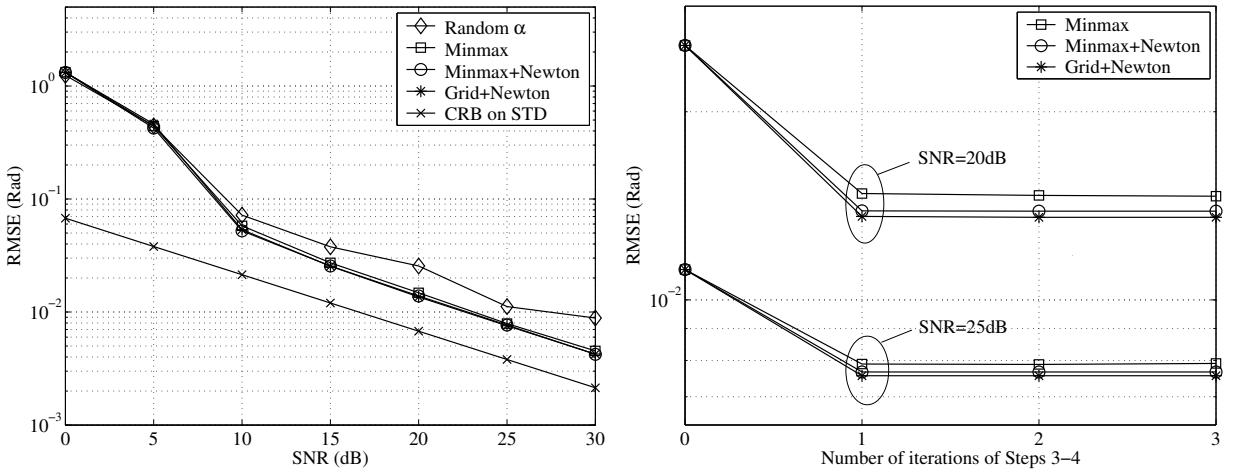
1. Given (14), follow (21)–(25) to obtain $\boldsymbol{U}_s$.
2. Randomly select $\boldsymbol{\alpha}$ subject to $|\alpha_n| = \sqrt{\frac{1}{N}}$, $n = 1, \ldots, N$, compute $\{\widehat{\omega}_{f,n}\}_{n=1}^{N}$, $f = 1, \ldots, F$, using (27)–(34).
3. Based on $\{\widehat{\omega}_{f,n}\}_{n=1}^{N}$, for $f = 1, \ldots, F$, obtain an updated $\boldsymbol{\alpha}_{\rm opt}$ by first solving (39) using SQP to get initials, then solving (37) using a Newton method.
4. Compute updated $\{\widehat{\omega}_{f,n}\}_{n=1}^{N}$, $1 \leq f \leq F$, with $\boldsymbol{\alpha}_{\rm opt}$ using (27)–(34).
5. Iterate Steps 3 and 4 until frequency estimates converge (typically one execution of Steps 3-4 is sufficient).

using the solution of (39) as $\boldsymbol{\alpha}_{\rm opt}$, referred to as "Minmax" here; (b) using the solution of (39) as initials, then solving (37) to obtain $\boldsymbol{\alpha}_{\rm opt}$, referred to as "Minmax+Newton", which is the proposed algorithm; and (c) solving (37) by grid search first then refine it using a Newton method, referred to as "Grid+Newton". These approaches are applied to estimate three 3-D frequency components from 3 snapshots of $6 \times 6 \times 6$ noisy data samples. Fig. 6 (a) depicts the Root Mean Square Error (RMSE) versus Signal-to-Noise Ratio (SNR). The RMSE is obtained by averaging over all frequencies after only one iteration of Steps 3-4, except for the case indicated with "Random α" where only Steps 1-2 are executed. The corresponding CRB on STD is also plotted.

It can be seen from Fig. 6 (a) that the three optimization methods provide similar performance, and all outperform the case with randomly chosen $\boldsymbol{\alpha}$. It turns out one iteration of Steps 3-4 is sufficient, as demonstrated by Fig. 6 (b), where we plot the RMSE of frequency estimates versus the number of iterations of Steps 3-4. Zero iteration corresponds to the case with only randomly chosen $\boldsymbol{\alpha}$. We observe that one execution of Steps 3 and 4 is sufficient as further iterations only provide negligible performance improvement if any. It can be seen that "Minmax+Newton" and "Grid+Newton" are comparable, and both are slightly better than "Minmax". Because "Grid+Newton" has a higher complexity than "Minmax+Newton", we choose "Minmax+Newton" with one iteration as the proposed algorithm in Table 3 for optimizing the weighting factors, which is the algorithm used in the simulations of Section B.6.3.

B.6.3 Simulation Results

In this section we present the Monte Carlo simulation results to demonstrate the performance (measured by RMSE) of the optimized eigenvector-based frequency estimation algorithm, which is also compared to other N-D frequency estimation algorithms as well as the associated CRB.



(a) (b)

Figure 6: (a) RMSE of different optimization methods versus SNR; (b) RMSE of different optimization methods versus the number of iterations of Steps 3-4.

2-D Identical Frequency Estimation from Single Snapshot

In the first experiment, the proposed algorithm, MEMP [37], Unitary ESPRIT [44] and 2-D ESPRIT [34] are applied to estimate three 2-D frequency components from a 20×20 noisy data set. The amplitudes, $c_f(1)$ for $f = 1, \ldots, F$, are set to be one for this case. The three frequency pairs are $(\omega_{1,1}, \omega_{1,2}) = (0.55\pi, 0.20\pi)$, $(\omega_{2,1}, \omega_{2,2}) = (0.60\pi, 0.20\pi)$, and $(\omega_{3,1}, \omega_{3,2}) = (0.60\pi, 0.25\pi)$. Notice that there are identical frequencies in both dimensions, which is a case that the IMDF algorithm fails to deal with. Fig. 7 depicts the performance comparison. In Fig. 7 (a), we plot the RMSE of various algorithms and the average CRB on STD in the two dimensions. The RMSE results are averaged over all frequencies and obtained through 1000 realizations. In Fig. 7 (a), "Proposed algorithm" refers to the one with one iteration of Steps 3-4 using "Minmax+Newton". For our proposed algorithm, the smoothing parameters $(\{K_n\}_{n=1}^2, \{L_n\}_{n=1}^2)$ are chosen such that the identifiability bound can be achieved. The parameters for other algorithms are chosen according to their respective references. As shown in Fig. 7 (a), the proposed algorithm offers comparable performance as that of Unitary ESPRIT, and outperforms 2-D ESPRIT and MEMP.

In Fig. 7 (b), we compare the optimized weighting factors to randomly chosen ones, where "Random α" means zero iteration of Steps 3-4 in Table 3. The theoretic RMSE is obtained with an optimized $\boldsymbol{\alpha}$ by solving (37) using the true frequencies, which serves as a benchmark since in our algorithm $\boldsymbol{\alpha}$ is optimized when the true frequencies are unknown. It is clear that the proposed algorithm significantly outperforms the one with random weighting factors, and the simulated RMSE of the proposed algorithm matches well to the theoretic RMSE for moderate to high SNR.

2-D Close Frequency Estimation from Multiple Snapshots

In the second experiment, the proposed algorithm, Unitary ESPRIT and RARE are applied to estimate three 2-D frequency components from 10 snapshots of noisy data, each of size 12×12, as

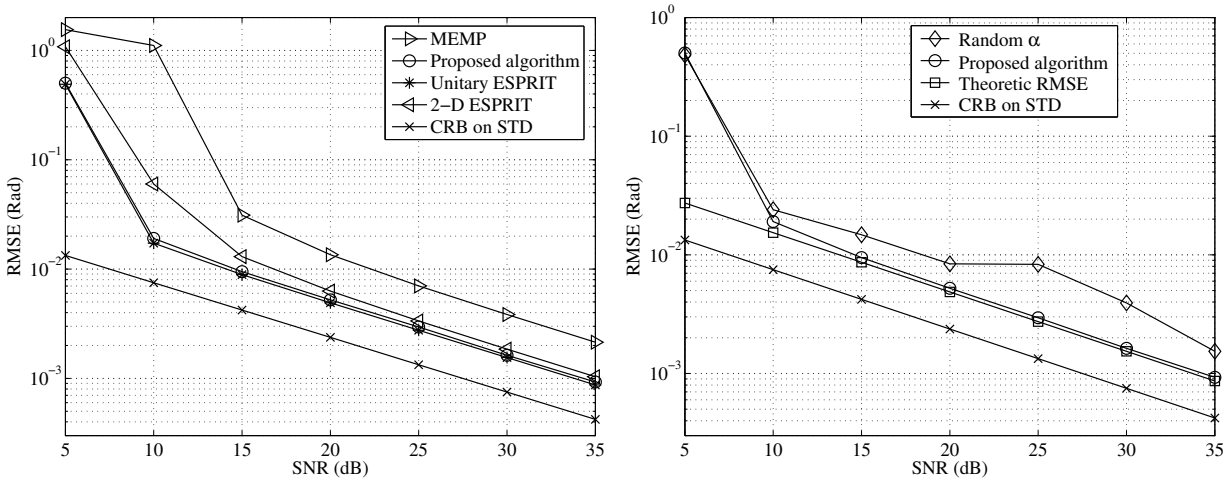


Figure 7: (a) Comparison of different algorithms for 2-D frequency estimation from single snapshot; (b) Comparison of optimized $\boldsymbol{\alpha}$ and randomly chosen $\boldsymbol{\alpha}$.

given in (14). The amplitudes, $c_f(t)$, for $f = 1, \ldots, F$ and $t = 1, \ldots, T$, are drawn from a complex Gaussian distribution. The three frequency pairs are $(\omega_{1,1}, \omega_{1,2}) = (0.72\pi, 0.62\pi)$, $(\omega_{2,1}, \omega_{2,2}) = (0.74\pi, 0.58\pi)$, and $(\omega_{3,1}, \omega_{3,2}) = (0.76\pi, 0.60\pi)$. Notice that frequencies are close to each other in both dimensions. Fig. 8 depicts the simulated RMSE of various algorithms, along with the corresponding CRB and the theoretic RMSE of the proposed algorithm. Multidimensional data smoothing is also performed for the Unitary ESPRIT algorithm and the RARE algorithm. It can be seen from Fig. 8 (a) that the proposed algorithm offers competitive performance when compared with the Unitary ESPRIT and RARE algorithms. Note that the proposed algorithm has lower complexity than these two algorithms, because the proposed algorithm does not require a frequency pairing step (the Unitary ESPRIT algorithm achieves automatic frequency pairing through iterative joint diagonalization). Fig. 8 (b) confirms again that optimized weighting factors outperform randomly chosen weighting factors, and the simulated RMSE of the proposed algorithm matches its theoretic RMSE at high SNR.

3-D Identical Frequency Estimation from Multiple Snapshots

In the third experiment, the proposed algorithm and the Unitary ESPRIT algorithm are applied to estimate three 3-D frequency components from 10 snapshots of $6\times6\times6$ noisy data samples. There are identical frequencies in all dimensions. The amplitudes are drawn from a complex Gaussian distribution. Fig. 9 shows the performance comparisons. From Fig. 9, we notice the proposed algorithm also offer competitive performance in 3-D frequencies estimation compared to the Unitary ESPRIT algorithm. Because the pairing strategy of RARE algorithm is not applicable when all three dimensions have identical frequencies, we do not include RARE in this experiment. Notice that the simulated RMSE of the proposed algorithm matches its theoretic RMSE for moderate to high SNR range.

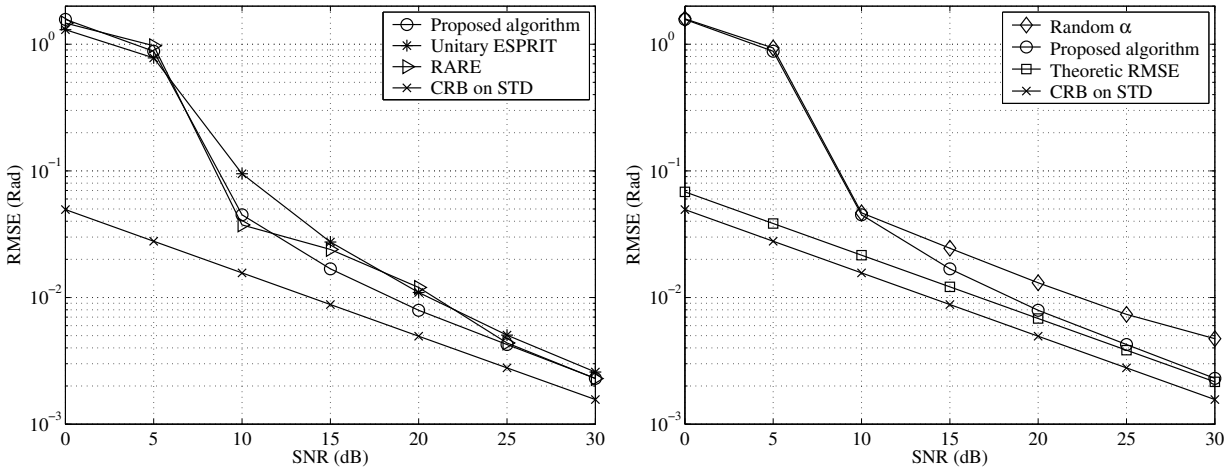


Figure 8: (a) Comparison of different algorithms for 2-D frequency estimation from multiple snapshots; (b) Comparison of optimized $\boldsymbol{\alpha}$ and randomly chosen $\boldsymbol{\alpha}$.

B.7 Conclusions

In this project we have proposed a signal processing framework for co-channel interference mitigation when multiple FH network coexist. The EM algorithm jointly estimates hop timing and frequency estimation of multiple FH signals, with unknown hop sequences and possible bandwidth mismatch. A simple initialization step based on the data spectrogram gives initial estimates of hop timing for the EM algorithm. Simulation results have shown that the EM algorithm is capable of obtaining the operation characteristic of noncooperative FH emitters. When there is a transmitter-receiver bandwidth mismatch in multiple FH networks, the model order changes evenif the number of active emitters does not vary. We have designed a low-complexity approach for model order variation detection based on the trace of the covariance matrix of the received signal. Simulation results demonstrate that the model order variation detection approach outperforms those based on sliding widow SVD or MDL.

As multidimensional frequency estimation plays an important role in collision resolution when multiple FH networks coexist, we have also proposed an eigenvector-based algorithm for N-D frequency estimation from multiple data snapshots. We have analytically quantified the identifiability (in the noiseless case) and the performance (in the noisy case) of the proposed algorithm. It is shown that our algorithm offers the most relaxed statistical identifiability bound to date. It remains operational when there exist identical frequencies in one or more dimensions, due to the adoption of weighting factors. Furthermore, a low-complexity (one iteration) approach is developed to optimize the weighting factors by minimizing MSEs of frequency estimates. Simulation results show that the proposed algorithm offers better or competitive performance when compared to existing algebraic approaches for N-D frequency estimation, but at lower complexity since frequency estimates are automatically paired without multiple iterations of joint diagonalization (as in the Unitary ESPRIT) or requiring an extra pairing step (as in MEMP or RARE). It is shown that the optimized weighting factors significantly outperform randomly chosen weighting factors.

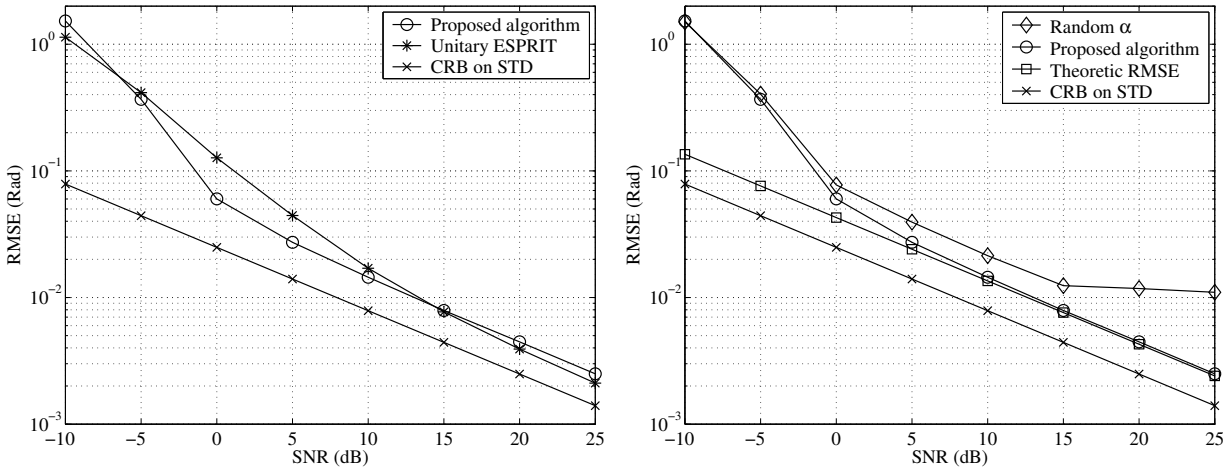


Figure 9: (a) Comparison of different algorithms for 3-D frequency estimation from single snapshot; (b) Comparison of optimized $\boldsymbol{\alpha}$ and randomly chosen $\boldsymbol{\alpha}$.

B.8 Bibliography

[1] D. J. Torrieri, "Mobile frequency-hopping CDMA systems," *IEEE Trans. Communications*, vol. 48, no. 8, pp. 1318–1327, Aug. 2000.

[2] Bluetooth Special Interest Group, Specification of the Bluetooth System, Version2.0+EDR, Nov. 2004.

[3] A. Conti, D. Darddari, G. Pasolini, and O. Andrisano, "Bluetooth and IEEE 802.11b coexistence: analytical performance evaluation in fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 2, pp. 259–269, Feb. 2003.

[4] I. Howitt, "Bluetooth performance in the presence of 802.11b WLAN," *IEEE Trans. Vehiclar Technology*, vol. 51, no. 6, pp. 1640–1651, Nov. 2002.

[5] I. Howitt, "Mutual interference between independent Bluetooth piconets," *IEEE Trans. Vehiclar Technology*, vol. 52, no. 3, pp. 708–718, May 2003.

[6] T.-Y. Lin, Y.-K. Liu, and Y.-C. Tseng, "An improved packet collision analysis for multi-Bluetooth piconets considering frequency-hopping guard time effect," *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 2087–2094, Dec. 2004.

[7] N. Golmie and F. Mouveaux, "Interference in the 2.4 GHz ISM band: impact on the Bluetooth access control performance," *Proc. IEEE Intl. Conf. on Communications*, vol. 8, pp. 2540–2545, Jun. 2001.

[8] J. Dunlop and N. Amanquah, "High capacity hotspots based on bluetooth technology," *IEE Proceedings-Communications*, vol. 152, no. 5, pp. 521–527, Oct. 2005.

[9] E. C. Arvelo, "Open-loop power control based on estimations of packet error rate in a Bluetooth radio," *Wireless Communications and Networking*, vol. 3, no. 5, pp. 1465–1469, Mar. 2003.

[10] N. Golmie and N. Chevrollier, "Techniques to improve Bluetooth performance in interference environments," *Proc. Military Communications Conference*, vol. 1, pp. 581–585, Oct. 2001.

[11] N. Golmie, N. Chevrollier, and O. Rebala, "Bluetooth and WLAN coexistence: challenges and solutions," *IEEE Wireless Communications*, vol. 10, no. 6, pp. 22–29, Dec. 2003.

[12] C. F. Chiasserini and R. R. Rao, "Coexistence mechanisms for interference mitigation in the 2.4-GHz ISM band," *IEEE Trans. Wireless Communications*, vol. 2, no. 5, pp. 964–975, Sep. 2003.

[13] C. D. M. Cordeiro, S. Abhyankar, R. Toshiwal, and D. P. Agrawal, "A novel architecture and coexistence method to provide global access to/from Bluetooth WPANs by IEEE 802.11 WLANs," *Proc. IEEE Int. Conf. on Performance, Computing, and Comm.*, pp. 23–30, Apr. 2003.

[14] B. Zhen, Y. Kim, and K. Jang, "The analysis of coexistence mechanisms of Bluetooth," *Proc. IEEE 55th Vehicular Technology Conference*, vol. 1, pp. 419–423, May 2002.

[15] IEEE Std. 802.15.2, "IEEE recommended practice for information technology – Part 15.2: coexistence of wireless personal area networks with other wireless devices operating in the unlicensed frequency bands," 2003, available at http://standards.ieee.org/getieee802/802.15.html.

[16] A. Willig, K. Matheus, and A. Wolisz, "Wireless technology in industrial networks," *Proc. of the IEEE*, vol. 93, no. 6, pp. 1130–1151, Jun. 2005.

[17] K. Yu-Kwong and M. C.-H. Chek, "Design and evaluation of coexistence mechanisms for Bluetooth and IEEE 802.11b systems," *Proc. IEEE Intl. Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 3, pp. 1767–1771, Sep. 2004.

[18] A. Souloumiac, "Blind source detection and separation using second-order non-stationarity," *Proc. International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1912–1915, May 1995.

[19] D. J. Torrieri and K. Bakhru, "Performance of recursive suppression algorithm in nonstationary environments," *IEEE Trans. Antennas and Propagation*, vol. 43, no. 3, pp. 299–307, Mar. 1995.

[20] U.-C.G. Fiebig, "An algorithm for joint detection in fast frequency hopping system," *Proc. International Conference on Communications*, vol. 1, pp. 90–95, Jun. 1996.

[21] T. Mabuchi, R. Kohno, and H. Imai, "Multiuser detection scheme based on canceling cochannel interference for MFSK/FH-SSMA system," *IEEE J. Sel. Areas of Communications*, vol. 12, no. 4, pp. 593–604, May 1994.

[22] L.-L. Yang and L. Hanzo, "Blind joint soft-detection assisted slow frequency-hopping multicarrier DS-CDMA," *IEEE Trans. Communications*, vol. 48, no. 9, pp. 1520–1529, Sep. 2000.

[23] L. Aydin and A. Polydoros, "Hop-timing estimation for FH signals using a coarsely channelized receiver," *IEEE Trans. Communications*, vol. 44, no. 4, pp. 516–526, Apr. 1996.

[24] M. K. Simon, U. Cheng, L. Aydin, A. Polydoros, and B. K. Levitt, "Hop timing estimation for noncoherent frequency-hopped M-FSK intercept receivers," *IEEE Trans. Communications*, vol. 43, no. 2/3/4, pp. 1144–1154, Feb./Mar./Apr. 1995.

[25] X. Liu, N. D. Sidiropoulos, and A. Swami, "Blind high resolution localization and tracking of multiple frequency hopped signals," *IEEE Trans. Signal Processing*, vol. 50, no. 4, pp. 889–901, Apr. 2002.

[26] X. Liu, N. D. Sidiropoulos, and A. Swami, "Joint hop timing and DOA estimation for multiple noncoherent frequency hopped signals," *Proc. 2nd IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 164–168, Aug. 2002.

[27] C. C. Ko, W. Zhi, and F. Chin, "ML-Based frequency estimation and synchronization of frequency hopping signals," *IEEE Trans. Signal Processing*, vol. 53, no. 2, pp. 403–410, Feb. 2005.

[28] X. Liu and N. D. Sidiropoulos, "On constant modulus multidimensional harmonic retrieval," *IEEE Trans. Signal Proc.*, vol. 50, no. 9, pp. 2366–2368, Sep. 2002.

[29] K. Konstantinides and K. Yao, "Statistical analysis of effective singular values in matrix rank determination," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. 36, no. 5, pp. 757–763, May 1988.

[30] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automatic Control*, vol. 19, no. 6, pp. 716–723, Dec. 1974.

[31] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. 33, pp. 387–392, Apr. 1985.

[32] M. Wax and I. Ziskind, "Detection of the number of coherent signals by the MDL principe," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. 37, no. 8, pp. 1190–1196, Aug. 1989.

[33] J. Li and X. Liu, "A collision resolution technique for robust coexistence of multiple Bluetooth piconets," *Proc. the 64th IEEE Vehicular Technology Conference*, Montreal, Quebec, Canada, Sept. 25-28, 2006.

[34] S. Rouquette and M. Najim, "Estimation of frequencies and damping factors by two-dimensional ESPRIT type methods," *IEEE Trans. Signal Processing*, vol. 49, no. 1, pp. 237–245, Jan. 2001.

[35] M. D. Zoltowski, M. Haardt, and C. P. Mathews, "Closed-form 2-D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT," *IEEE Trans. Signal Processing*, vol. 44, no. 2, pp. 316–328, Feb. 1996.

[36] A. N. Lemma, A.-J. van der Veen, and E. F. Deprettere, "Analysis of joint angle-frequency estimation using ESPRIT," *IEEE Trans. Signal Processing*, vol. 51, no. 5, pp. 1264–1283, May 2003.

[37] Y. Hua, "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Trans. Signal Processing*, vol. 40, no. 9, pp. 2267–2280, Sep. 1992.

[38] X. Liu, N. D. Sidiropoulos, and T. Jiang, "Multidimensional harmonic retrieval with applications in MIMO wireless channel sounding," in A. B. Gershman and N. D. Sidiropoulos, editors, *Space-Time Processing for MIMO Communications*, Wiley, pp. 41-75, 2005.

[39] X. Liu, J. Li, and X. Ma, "An EM algorithm for blind hop timing estimation of multiple frequency hopped signals with bandwidth mismatch," *IEEE Trans. Vehicular Technology*, to appear, 2007.

[40] S. Valaee and P. Kabal, "An information theoretic approach to source enumeration in array signal processing," *IEEE Trans. Signal Processing*, vol. 52, no. 5, pp.1171–1178, May 2004.

[41] J. Li, X. Liu, and A. Swami, "Detection of model order variations in frequency hopped systems," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 680–684, New York City, NY, Jun. 2005.

[42] S. M. Kay, *Fundamentals of Statistical Signal Processing, vol. II, Detection Theory*, Prentice-Hall, 1998.

[43] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47–60, Nov. 1996.

[44] M. Haardt and J. A. Nossek, "Simultaneous Schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Processing*, vol. 46, no. 1, pp. 161–169, Jan. 1998.

[45] M. Pesavento, C. F. Mecklenbräuker, and J. F. Böhme, "Multi-dimensional rank reduction estimator for parametric MIMO channel models," *EURASIP Journal on Applied Signal Processing*, vol. 2004, pp. 1354–1363, Sept. 2004.

[46] J. Liu, X. Liu, and X. Ma, "Multidimensional frequency estimation with finite snapshots in the presence of identical frequencies," *IEEE Trans. Signal Processing*, submitted, May 2006.

[47] D. J. Hartfiel, *Matrix Theory and Applications with MATLAB*, CRC Press, 2001.

[48] J. Liu and X. Liu, "An eigenvector-based approach for multidimensional frequency estimation with improved identificability," *IEEE Trans. Signal Processing*, vol. 54, no. 12, pp. 4543–4556, Dec. 2006.

[49] J. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits and Systems*, vol. 25, no. 9, pp. 772–781, Sept. 1978.

[50] K. V. Mardia and P. Jupp, *Directional Statistics*, 2nd ed., Wiley, 2000.

[51] R. W. Freund and F. Jarre, "Solving the sum-of-ratios problem by an interior point method," *Journal of Global Optimization*, vol. 19, no. 1, pp. 83–102, Jan. 2001.